

1:	gb.ba:*
2:	gb.hbg:*
3:	gb.in:*
4:	gb.om:*
5:	gb.ov:*
6:	gb.pat:*
7:	gb.ph:*
8:	gb.pl:*
9:	gb.pri:*
10:	gb.ro:*
11:	gb.sts:*
12:	gb.syz:*
13:	gb.un:*
14:	gb.yi:*
15:	em.ba:*
16:	em.fun:*
17:	em.hum:*
18:	em.in:*
19:	em.mu:*
20:	em.om:*
21:	em.or:*
22:	em.ov:*
23:	em.pat:*
24:	em.ph:*
25:	em.pl:*
26:	em.ro:*
27:	em.sts:*
28:	em.un:*
29:	em.yi:*
30:	em.hbg.hum:*
31:	em.hbg.in:*
32:	em.hbg.om:*
33:	em.hbg.ov:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description

1	605	100.0	605	6	AX016311	AX016311	Sequence
2	605	100.0	622	8	AF05027	AF05027	Brassicaceae
3	297.4	49.2	818	6	AX016324	AX016324	Sequence
4	289.4	48.8	576	6	AX016322	AX016322	Sequence
5	171.2	28.3	110804	8	ATAC016629	ATAC016629	Arabiaceae
6	106.2	11.6	1718	6	AX016320	AX016320	Sequence
7	54	8.9	7218	6	166494	166494	Sequence
8	52.2	8.6	3666	1	AF319618	AF319618	Vibrionaceae
9	51.4	8.5	11752	2	AL596454	AL596454	Human DNA
10	49.4	8.2	15342	3	AEO04420	AEO04420	Plasmid
11	49.2	8.1	77835	3	PFMA01352-3	PFMA01352-3	Conjunction (4 of 4)
12	49	8.1	193488	30	AC018519	AC018519	Homosapiens
13	49	8.1	198509	2	AC108142	AC108142	Homosapiens
14	48.8	8.1	1324	6	AC016326	AC016326	Sequence
15	48.8	8.1	151498	9	AC099331	AC099331	Homosapiens
16	48.8	8.1	190748	9	AC010965	AC010965	Homosapiens
17	48.2	8.0	107770	2	AC098572	AC098572	Oryza sativa
18	48	7.9	122583	2	AC092435	AC092435	Homosapiens
19	48	7.9	141710	2	AC012348	AC012348	Homosapiens
20	47.8	7.9	81202	2	AC068000	AC068000	Homosapiens
21	47.8	7.9	124057	2	AC068000	AC068000	Homosapiens
22	47.8	7.9	124057	2	AP000770	AP000770	Homosapiens
23	47.8	7.9	132441	2	AC027776	AC027776	Homosapiens
24	47.8	7.9	168494	2	AL450083	AL450083	Human DNA
25	47.8	7.9	175706	2	AC013798	AC013798	Homosapiens
26	47.8	7.9	171747	9	AC007539	AC007539	Homosapiens
27	47.8	7.9	179753	9	AC096734	AC096734	Homosapiens
28	47.8	7.9	180851	9	AL365179	AL365179	Human DNA
29	47.6	7.9	55149	9	AL068682	AL068682	Human DNA
30	47.6	7.8	226	8	AY018821	AY018821	Oryza sativa
31	47.4	7.8	135997	2	AC105261	AC105261	Oryza sativa
32	47.4	7.8	146465	9	AC026273	AC026273	Homosapiens
33	47.4	7.8	156387	2	AC013593	AC013593	Homosapiens
34	47.4	7.8	157373	2	AL593856	AL593856	Human DNA
35	47.4	7.8	166093	2	AC106013	AC106013	Homosapiens
36	47.4	7.8	173794	9	AC023108	AC023108	Homosapiens
37	47.2	7.8	15951	1	AEO01073	AEO01073	Archaeoglobus
38	47.2	7.8	161835	2	AC104794	AC104794	Homosapiens
39	47	7.8	97160	9	CNS06C7	CNS06C7	Human chromosome
40	47	7.8	149278	2	AC011602	AC011602	Homosapiens
41	47	7.8	183082	9	AC019183	AC019183	Homosapiens
42	46.8	7.7	141854	2	AC021595	AC021595	Homosapiens
43	46.8	7.7	161551	2	AP004259	AP004259	Homosapiens
44	46.6	7.7	155712	9	AP004219	AP004219	Homosapiens
45	46.6	7.7	177730	9	AC087844	AC087844	Homosapiens

ALIGNMENTS

RESULT 1	AX016311	605 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	AX016311				
DEFINITION	Sequence 14 from Patent WO9949046.				
ACCESSION	AX016311				
VERSION	AX016311.1	GI:10041874			
KEYWORDS	.				
SOURCE	tape.				
ORGANISM	Brassica napus				

REFERENCE 1 (bases 1 to 605)
AUTHORS Roberts,J.A., Wyatt,P. and Whitlaw,C.
TITLE Signal transduction protein involved in plant deniscence
JOURNAL Patent: WO 9943046-A 14 30-SEP-1999;

ROBERTS JEREMY ALAN (GB); BIOGEMMA UK LTD (GB); WYATT PAUL (GB);
WHITELAW CATHERINE (GB)

```
FEATURES      Location/Qualifiers
source        1. .605
              /organism="Brassica napus"
              /db_xref="taxon:3708"
CDS           20. .430
```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07684.1"
/db_xref="GI:10041875"
/translation="WATKSMGDEIKKKIKKUNLVYDDPLNLIIEKTIKAGISQT
ANNEEAVIHRDGSSEFDLIMDKEMPERDGVSTYRKUREMEVSMIVGTSIADNE
ERRAFMEAGINHLAPLTKDKIIPILNOLMDA"
BASE COUNT 198 a 101 c 140 g 166 t
* ORIGIN

Query Match 100.0%; Score 605; DB 6; Length 605;
Best Local Similarity 100.0%; Pred. No. 9.8e-138;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCACGACGAGATGCAAGATGCGCAAAATCCATGCGAGATATCGAAAAATAAGAA 60
1 GGCACGACGAGATGCAAGATGCGCAAAATCCATGCGAGATATCGAAAAATAAGAA 60
61 GAACTAAACGCTGTGATCGTCGATGATGATCCATTAACCTTATATGAGAAGAT 120
61 GAACTAAACGCTGTGATCGTCGATGATGATCCATTAACCTTATATGAGAAGAT 120
121 CATCAAGCGATGCGGGGATTTTCACAGACGCAATTAAGGTGAGAGCGCATATCAT 180
121 CATCAAGCGATGCGGGGATTTTCACAGACGCAATTAAGGTGAGAGCGCATATCAT 180
181 CCACCGTACGCGGCGCTCATCTTTGACCTTATCTTATGATTAAGAAATGCCGAGAG 240
181 CCACCGTACGCGGCGCTCATCTTTGACCTTATCTTATGATTAAGAAATGCCGAGAG 240
181 CCACCGTACGCGGCGCTCATCTTTGACCTTATCTTATGATTAAGAAATGCCGAGAG 240
241 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 300
241 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 300
301 GGTGACTTACGCTGCTGCAATGAGAGAGCGGCGCTTCACTGGAAGCTGACTTAA 360
301 GGTGACTTACGCTGCTGCAATGAGAGAGCGGCGCTTCACTGGAAGCTGACTTAA 360
361 CCATTCCTTGGCAAAACCGTTAACCAAGGACAGATCATCTCTCATTAACCACTCAT 420
361 CCATTCCTTGGCAAAACCGTTAACCAAGGACAGATCATCTCTCATTAACCACTCAT 420
421 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 480
421 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 480
481 TATGTATGATAGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 540
481 TATGTATGATAGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 540
541 GATATATATCATGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 600
541 GATATATATCATGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 600
601 CTACC 605
601 CTACC 605

RESULT 2
AF057027 622 bp mRNA linear PLN 01-OCT-1998
LOCUS Brassica napus response regulator protein mRNA, complete cds.
DEFINITION AF057027
ACCESSION AF057027
VERSION AF057027.1 GI:3687687
KEYWORDS
SOURCE
ORGANISM
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica
REFERENCE
1 (bases 1 to 622)
WhiteLaw,C.A., Paul, W., Jenkins, E.S., Taylor, V.M. and Roberts, J.A.

TITLE A mRNA encoding a response regulator protein from Brassica napus is
up-regulated during pod development
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 622)
AUTHORS WhiteLaw, C.A.
TITLE Direct Submission
JOURNAL Submitted (02-Apr-1998) School of Biological Sciences, University
of Nottingham, Sutton Bonington Campus, Loughborough,
Leicestershire LE12 5RD, England U.K.
FEATURES
source
location/Qualifiers
1..622
/organism="Brassica napus"
/db_xref="taxon:3708"
20..430
/note="up-regulated during pod development"
/codon_start=1
/product="response regulator protein"
/protein_id="CAC07684.1"
/db_xref="GI:3687688"
/translation="WATKSMGDEIKKKIKKUNLVYDDPLNLIIEKTIKAGISQT
ANNEEAVIHRDGSSEFDLIMDKEMPERDGVSTYRKUREMEVSMIVGTSIADNE
ERRAFMEAGINHLAPLTKDKIIPILNOLMDA"
BASE COUNT 215 a 101 c 140 g 166 t
ORIGIN

Query Match 100.0%; Score 605; DB 8; Length 622;
Best Local Similarity 100.0%; Pred. No. 9.8e-138;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCACGACGAGATGCAAGATGCGCAAAATCCATGCGAGATATCGAAAAATAAGAA 60
1 GGCACGACGAGATGCAAGATGCGCAAAATCCATGCGAGATATCGAAAAATAAGAA 60
61 GAACTAAACGCTGTGATCGTCGATGATGATCCATTAACCTTATATGAGAAGAT 120
61 GAACTAAACGCTGTGATCGTCGATGATGATCCATTAACCTTATATGAGAAGAT 120
121 CATCAAGCGATGCGGGGATTTTCACAGACGCAATTAAGGTGAGAGCGCATATCAT 180
121 CATCAAGCGATGCGGGGATTTTCACAGACGCAATTAAGGTGAGAGCGCATATCAT 180
181 CCACCGTACGCGGCGCTCATCTTTGACCTTATCTTATGATTAAGAAATGCCGAGAG 240
181 CCACCGTACGCGGCGCTCATCTTTGACCTTATCTTATGATTAAGAAATGCCGAGAG 240
181 CCACCGTACGCGGCGCTCATCTTTGACCTTATCTTATGATTAAGAAATGCCGAGAG 240
241 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 300
241 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 300
301 GGTGACTTACGCTGCTGCAATGAGAGAGCGGCGCTTCACTGGAAGCTGACTTAA 360
301 GGTGACTTACGCTGCTGCAATGAGAGAGCGGCGCTTCACTGGAAGCTGACTTAA 360
361 CCATTCCTTGGCAAAACCGTTAACCAAGGACAGATCATCTCTCATTAACCACTCAT 420
361 CCATTCCTTGGCAAAACCGTTAACCAAGGACAGATCATCTCTCATTAACCACTCAT 420
421 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 480
421 GGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAATGATCAATGATGTTGG 480
481 TATGTATGATAGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 540
481 TATGTATGATAGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 540
541 GATATATATCATGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 600
541 GATATATATCATGATGATGCTGATGCTGTTTGAATTTAGAGTCTTATGCGCGT 600
601 CTACC 605
601 CTACC 605

RESULT 3	AX016324	AX016324	818 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	AX016324					
DEFINITION	Sequence 27 from Patent WO9949046.					
ACCESSION	AX016324					
VERSION	AX016324.1	GI:10041879				
KEYWORDS						
SOURCE	thale cress.					
ORGANISM	Arabidopsis thaliana					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 818)					
AUTHORS	Roberts,J.A., Wyatt,P. and Whitelaw,C.					
TITLE	Signal transduction protein involved in plant dehiscence					
JOURNAL	Plant J. MO 9949046-A 27 30-SEP-1999;					
FEATURES	ROBERTS JEREMY ALAN (GB); BIOEMMA UK LTD (GB); WYATT PAUL (GB); WHITELAW CATHERINE (GB)					
Source	Location/Qualifiers					
	1..818					
	/organism="Arabidopsis thaliana"					
	/db_xref="taxon:3702"					
	180..608					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAC07686.1"					
	/db_xref="GI:10041880"					
	/translation="MATKSTGTEKTSIEVKKLLINVLIVDDPELNRLEHMIKTTGGISOTAKNGEAVLHEDGASFDLILMKKEPREDGVSTIKXLREMGKSMIVGYTVAADDEERKAFMEGDLNHCHLEKPLTFAKIFLPSILFPA"					
BASE COUNT	287 a 148 c 159 g 222 t					2 others
ORIGIN						
Query Match	49.2%; Score 297.4; DB 6; Length 818;					
Best Local Similarity	71.6%; Pred. No. 1.7e-62;					
Matches 449; Conservative 0; Mismatches 138; Indels 40; Gaps 3;						
OY	12 AATGAGAGTGGGACAAATAATCCATGGGAGATCGAGAAATATAGAAACTA----- 67					
DB	172 AATGAGAGTGGGACAAATAATCCATGGGAGATCGAGAAATATAGAAACTA----- 67					
OY	68 -----AAGGTGTGATCGTGCATGATGATGATCATCACTAAACCTTATATTCATG 113					
DB	232 AGAAGAACTAATCAACGCTGTGATCGTGCATGATGATGATCATCACTAAACCTTATATTCATG 113					
OY	114 AGAAGATCATCAAGGATTTGGGGGATTTCAAGACAGCAATTAACGGTGGAGAGCGAG 173					
DB	292 AGATGATCATCAAGGATTTGGGGGATTTCAAGACAGCAATTAACGGTGGAGAGCGAG 173					
OY	174 TTAATCATCACCGTGAAGCGCGCTATCTTTTGAACCTTATATGATGATGAAGAAATGC 233					
DB	352 TGATCTCCACCGCTGAAGCGCGCTATCTTTTGAACCTTATATGATGATGAAGAAATGC 233					
OY	234 CCGAGAGGATGGTGTGGACACATCAAGAACTTAAGAAATGGAAGTGAATCATATGA 293					
DB	412 CTGAGAGGATGGTGTGGACATTAAGAACTTAAGAAATGGAAGTGAATCATATGA 293					
OY	294 TTGTTGGGTGACTTGCTGACAAATGAAGAGAGACCGAGGCTTTTATGGAAGCTG 353					
DB	472 TCGTTGGGTGACTTGCTGACAAATGAAGAGAGACCGAGGCTTTTATGGAAGCTG 353					
OY	354 GACTTAACCATTCCTTGGCAAAACCGTTAACCAGAGCAAGATATATCCCTCTATTAACG 413					
DB	532 GGGTCAACCATTCCTTGGCAAAACCGTTAACCAGAGCAAGATATATCCCTCTATTAACG 413					
OY	414 AACTCATGATGCTTATGATGATATATATTATTA-----TTATGAAACACATCATATA 467					
DB	592 AACTCTCATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 467					
OY	468 ACGCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 521					
DB	652 TCACCTTAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711					

Oy	522	-----AGGGTCTTTACGCGCGTGATATAATATCATGTAGTGTGCTTTAAG	571
Db	712	TCTATGCTTGATTATTATTTCCTATCGCCGAGGAAPAAATATCATGCAAGTCATTTCTTTGG	771
Oy	572	CTTATAAATATTTAATAGAAGTTTC	598
Db	772	CTATATAAATATTTAATAGAAGTTTTC	798
RESULT 4			
LOCUS	AX016322	576 bp	DNA
DEFINITION	Sequence 25 from Patent WO9494046.		linear
ACCESION	AX016322		PAT 07-SEP-2000
VERSION	AX016322.1	GI:10041878	
KEYWORDS			
SOURCE	Rape.		
ORGANISM	Brassica napus		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
REFERENCE	1 (bases 1 to 576)		
AUTHORS	Roberts,J.A., Wyatt,P. and Whitelaw,C.		
TITLE	Signal transduction protein involved in plant dehiscence		
JOURNAL	Patent: WO 9494046-A 25 30-SEP-1995;		
	ROBERTS JEREMY ALAN (GB); BIOGENMA UK LTD (GB); WYATT PAUL (GB);		
	WHITELAW CATHERINE (GB)		
FEATURES	Location/Qualifiers		
source	1..576		
	/organism="Brassica napus"		
	/db_xref="taxon:3708"		
BASE COUNT	188 a 102 c 122 g 161 t	3 others	
ORIGIN			
Query Match	47.8%; Score 289.4; DB 6; Length 576;		
Best Local Similarity	75.4%; Pred.No.1.6e-60;		
Matches 416; Conservative 0; Mismatches 94; Indels 42; Gaps 3;			
Oy	78	TCGTGATGATGATCCACTTAACCTTATTAATTCATGAGAAGATCATCAAGCATTGGCG	137
Db	1	TCGCATGTGTGATCCCTGAATACGTCAACCTTCACGAGATTATCATCAATCATATCGGTG	60
Oy	138	GATTTACAGAAGAGGATTAACGGTGAAGAGCGCGATATCATCCACCGTAGCGGGCT	197
Db	61	GAATTCACAGACAGCTAAGAACGGTGAAGAGCGAGTGAACATCCACCGGACGGCATAG	120
Oy	198	CATCTTTGACCTTATCTCATATGATATAAGAATGCCGAGAGGATGTGTTTGACAA	257
Db	121	CATCTTTGACCTTATCTCATATGATATAAGAATGCCGAGAGGATGTGACTTTCGGCAA	180
Oy	258	CTAAGAAAGCTAAAGAAAATGSAAGAGATCTAATGATTGTGGGGAGCTCTACTGGCTG	317
Db	181	CTAAGAAAGCTAAAGAAAATGSAAGAGATCTAATGATTGTGGGGAGCTCTACTGGCTG	240
Oy	318	ACAATGAGAGGAGCGGCGAGGCTTTCATGAGAGCTGACCTTAACCATTTGCTGGCAAAAC	377
Db	241	ACAATGAGAGGAGCAACTTAAGGCTTTCATGAGAGGAGCTGACCTTAACCATTTGCTGGCAAAAC	300
Oy	378	CGTAACCAAGSACAAGATCATCCCTGTCATATCAACAACATCATGATGGCTGATGGA---	434
Db	301	CCTTAAGCAAGSCAAGATCTCCCTCTCATCAACAACATCTCATGATGATGCTTATGATGAG	360
Oy	435	-----TATATATTTTATATTAATGAGAAACACATATAAGCT	471
Db	361	ATGATTTGCGCCACATATATCTCATTAATACAAATATGAAAAAGCACATATATAACT	420
Oy	472	-----CTAAGTGTATATGATGATGATGACTTGATGATGTGTGTTTAGAATTAG	524
Db	421	CATACACCTGTGTGTATGATATATATCTATCCGATGTGTGTTTTTAGGTTTAT	480
Oy	525	GTT-----CTTATCGTCGATATATAATCATGTAAGTGTGTGCTTTAAAGCTTA	575

[illegible]

Db 2947 AATATCAGAGAAAATTAATAATCTCAAGGTATTGCTTACACAGCAATGCCCT 3006

QY 322 TGAAGAGAGCGAGCGGCTTTCATGAGAGCTGACATTACCATGTGTTGGCAAAACGTT 381

Db 3007 TCAAAAAGCAGCATGATACCTTTTAAAGAGAGGCAACCTATGCTTAACTACCAT 3066

QY 382 AACCAAGACAGATCATCCCTCTCATTAACCAACATCATGATGC 426

Db 3067 ACAAGAAATGATTCTTCTGGCGCATCAACATATCAAGATGC 3111

RESULT 9
AL596454 111752 bp DNA linear PRI 18-DEC-2001
AL596454
LOCUS Human DNA sequence from clone Rpl1-17718 on chromosome 9, complete
DEFINITION sequence.
ACCESSION AL596454.8 GI:17939740
VERSION HTG.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 111752)
TITLE Direct Submission
JOURNAL Submitted (18-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On Dec 19, 2001 this sequence version replaced gi:11784505.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: EM, EMBL; S, S;
SWISSPROT, Tr., TrEMBL; WP., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/BSG/Chr9
Rpl1-17718 is from the library RPl1-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone
Rpl1-17718. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone Rpl1-515013 is at 109753 in this
sequence.

FEATURES
Location/Qualifiers
Source
1..111752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="Rpl1-17718"
/clone_lib="RPl1-11.1"
BASE COUNT 34576 a 19899 c 20215 g 37062 t
ORIGIN
Query Match 8.5%; Score 51.4; DB 9; Length 111752;
Best Local Similarity 56.1%; Pred. No. 0.021;

Matches 97; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 427 TTGATGAGATATATTTATTTATATATGAAACACATATTAACGCTTAAGTGTATATA 486

Db 80799 TTGATGAGATATATTTATTTATATATGAAACACATATTAACGCTTAAGTGTATATA 486

QY 487 TGATGATATCTGCAGTGTGTGTTTGAATTTAGAGGTTCTTATCGTCGATATATA 546

Db 80859 TCGTGATATCATGTGTGTGTGTGTTTTCATATGATATTTCTTTCTGTAGTGTGTATCT 80858

QY 547 TAATCATGTAAGTTGTTGCTTTAAGCTTTAAATATTTAAATAGGTTTCC 599

Db 80919 TAAATTTGTAATTTATTTATGCACTGAGGTAAATAATTTCTGTATATCAGCTCC 80971

RESULT 10
AE001420/c 15342 bp DNA linear INV 06-NOV-1998
AE001420/c
LOCUS Plasmodium falciparum chromosome 2, section 57 of 73 of the
DEFINITION complete sequence.
ACCESSION AE001420 AE001362
VERSION AE001420.1 GI:3845287
KEYWORDS
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS 1 (bases 1 to 15342)
Gardner, M.J., Tetteh, H., Carucci, D.J., Cummings, L.M., Aravind, L.,
Koonin, E.V., Shallow, S., Mason, T., Yu, K., Fujii, C., Pederson, J.,
Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Pertea, M.,
Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, O.,
Smith, H.O., Fraser, C.M., Hoffman, S.L. et al.
Chromosome 2 sequence of the human malaria parasite Plasmodium
falciparum
Science 282 (3391), 1126-1132 (1998)

JOURNAL MEDLINE
99021743
REMARK Erratum: [[published erratum appears in Science 1998 Dec
4;282(5395):1827]]
REFERENCE 2 (bases 1 to 15342)
Gardner, M.J.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-1998) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20814, USA

FEATURES
Location/Qualifiers
Source
1..15342
/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="2"
2946..3269
/gene="PFB0830w"
2946..3269
/gene="PFB0830w"
/note="Identified by sequence similarity; putative"
/codon_start=1
/product="ribosomal protein S26"
/protein_id="AAC71955.1"
/db_xref="GI:3845288"
/translation="MPKRBNGRGSRKHNRGVNPLRGSNCGRCVPKDKAKRIRIRNI
VPSNRQDKKSVSYSTQLPKLYIKQCYCSALHSFVVRBQRARRKETAKEY
NPSQL"
complement(4112..5671)
/gene="PFB0835c"
complement(4112..5671)
/gene="PFB0835c"
/note="predicted by Glimmer"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71956.1"
/db_xref="GI:3845289"
/translation="MKRYINKSKALLLRKMSSTVKNMSKSNQLEIFMALEKETSLLQ
KEKLYIEPEVENICSLNETSCPLTIDVYFTLWYKSRIDITVYELKRVYPCGSGGTNNMG
LNNNDKTSSTDKGFIINKKEGEERKKEETNMNENKRVDFDSFVQAINIVLDFEK
VFYKHIILIDSLAADVYRHSIGVYRNRKYNRVYLVFNLLKTYEKNNVHHHESFA

```
LHDSYPLINCYIVNSGTSYLLKLPFOQNNLVEFYITTFVAVNCLTFKPOQVYVPL
NSMONEKQIKSLCSHIDKQVSLFENKREYKSLTHLHDEMLISNFKKYOONLGYI
KINNEYTYINDEONELVONSNDNLTIPSSILNDEDFSHAKGVQYACKRDVQSNH
FISSEPKNIHNEFEGHLYLSPCHAMVYIKSLIPRIIMQISIOYILIPVPSVSL
YKGYRYSDSHSIOVHYATYCFDKKRAYITLTICDMSFENSDISFVQSOEERKLY
MLGIRSLDEP
7534.8526
/gene="PF0840w"
7534.8526
/gene="PF0840w"
7534.8526
/gene="PF0840w"
/note="identified by sequence similarity; putative"
/codon_start=1
/product="replication factor C, 40 kDa subunit
(replication activator)"
/protein_id="AAC71957.1"
/db_xref="GI:3845290"
/translation="MENIPWYEKYRPRKLDIVHONNAVMKLEKVTYKMPHLFPHG
PPGKTSATNALAELEFGENISERYELNADSDGINVVERIKAYRISIKNKI
HSETRKVPKRIYVDEADMETEDASQARILEIYSNTPELICYTHIKSDPIT
SRCSYRPOSIPINIKKELIYICQNDIYDADLEKITEGDRRAVSLQDLS
CINRITLNSVLDSGLPSDNIVKIIDACKMDKLVEYTDIIEGDAVAYIKRS
FNNFVNTYEDSLKYQILLESRDYRLHGCATYIQLSPASSVSHLSNV"
9688.10611
/gene="PF0845w"
9688.10611
/gene="PF0845w"
/note="predicted by Glimmer"
/codon_start=1
/product="predicted integral membrane protein"
/protein_id="AAC71958.1"
/db_xref="GI:3845291"
/translation="MDSYIEKRSNVLNKOYDIKIONEYDELTSYSIDKYSSEDM
LNKYEKTSIDEDYMYOKINDNDNSERDNCINYEYIEMTEKREKREKMTL
ESINCYDFENDPFLIKVPEKISNSMKNHNIKLEFLVLTYSFTFTFR
KYTNLPEKHNLFLEKVENIKMHRKNSVLEFYDRIRYIVNKFEDVPO
FLCLLTLLNGKDFIFENMLYSIIYFVNAIYHGLNPLMGLNLFSTYLLDYI
NKKRKNLYIYIYICMYLFLCIYINFTI"
complement(11963..14386)
/gene="PF0850c"
complement(11963..14386)
/gene="PF0850c"
complement(11963..14386)
/note="predicted by Glimmer"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71959.1"
/db_xref="GI:3845292"
/translation="MAVESKPNSSKREKNEENDLINKDQSNKINKENFEAYEKGI
NESGMSNDINNKOEKREKKNTHKRYNINHTDNRKNNQDIPKPEI
EDNITINKNPTNILDSEYEGENRNDINNNNNNNINNNNNKNNNNNN
KKTITLKRNDIKDGINNENITTLNKNNDKNNNNNDNRNNNNNNNNNNNN
CCSEKTLQREKEYNIRARLESNFKQKNOVQTEQNNLNTYLNNNIINNINNDN
OYAYINNHYHYNNSYHIERQNNIPICINNHAPIEKLNPPYHDIHAYIYV
YSTOKNNMKTKOIGHYINNEDNNNNNNNNNNNNNNNNNNNNNNNNNNNN
DNYKTKNNFQGTNNFNOGTNNFCTNNFNNNAKNNIKINNTNNKNIHELNNHSIY
NVEYENKNIYDANGNINNINISYTLCKNNNNINENIHMSPIQCHNNTEKYNNDT
PNEPTNNKKKREKRIHENNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
VKNTEONLQKRNKQKQSKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
VDNKLKSSHADNSNEVTKGRKNNKNNKNNIINNIVANNVNNNNNNNNNN
SMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NVMNNNNLTSNENYNNIINISYEPFEYENLMDLYCRDISLYEKYRDGNDIQNHK
RYDIFPPL"
BASE COUNT 6057 a 1268 c 1441 g 6576 t
ORIGIN
Query Match 8.2%; Score 49.4; DB 3; Length 15342;
Best Local Similarity 52.6%; Pred. No. 0.07;
Matches 130; Conservative 0; Mismatches 116; Indels 1; Gaps 1;
QY 333 GCAGGCTTTCATGAGAGTGAACCATTCCTGGCAAAACCGTTACCAAGGACA 392
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7604 GCATTTATTTTGAAGACATGTCACCAACCTTTTGGTCGTTCTTCAACCAAGGA 7545
QY 393 AGATCATCCCTCTCATTAACCAACTGATGCTGATGATATATATTTATATTATG 452
```

```
DB 7544 AATATTTTCATCTTAAAAATTA- AAATTAATAATATATATATATATATATA 7486
QY 453 GAAACACACATATATACCTCTAAGTGTATGTATGATATCTTGATGTGTGT 512
DB 7485 TATATATATATATATATATATATATATGATATGATATATATATATATAT 7426
QY 513 TTAGATTTAGGTTCTTTATCGCCGATATATATATATATATATATATATAGC 572
DB 7425 TATATATATATATATATATATATATATATATATATATATATATATATAT 7366
QY 573 TTATATA 579
DB 7365 ATATATA 7359
RESULT 11
PFMAL13P2_3/c
wpcoment
Sequence split into 4 fragments LOCUS PFMAL13P2 Accession AL049185
Fragment Name Begin End
PFMAL13P2_0 1 110000
PFMAL13P2_1 100001 210000
PFMAL13P2_2 200001 310000
PFMAL13P2_3 300001 377835
Continuation (4 of 4) of PFMAL13P2 from base 300001 (AL049185 Plasmodium falciparum)
Query Match 8.1%; Score 49.2; DB 2; Length 77835;
Best Local Similarity 51.4%; Pred. No. 0.074;
Matches 114; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 380 TTAACCAAGGACAATATATCCCTCTCATTAACCACTGATGATGATATAT 439
DB 56551 TGAAGCGAGATACGACGACGACGACGCTTAAACCCCTGGTCCAGAAATTTAA 56492
QY 440 ATTTATATATATAGAACACATATATATATATATATATATATATATATAT 499
DB 56491 AATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 56432
QY 500 GCATGCTGTCTTTAGATTTAGGTTCTTATCGCCGATATATATATATATAT 559
DB 56431 GATATATATATATATATATATATATATATATATATATATATATATAT 56372
QY 560 TGTGCTTAACTTAAATATTTAAATTAAGGTTTCCTC 601
DB 56371 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 56330
RESULT 12
AC018519
ID AC018519 standard; DNA; HTG; 193488 BP.
XX AC018519;
XX AC
XX AC018519.4
XX SV
XX 15-DEC-1999 (Rel. 62, Created)
DT 28-MAY-2000 (Rel. 63, Last updated, Version 3)
XX Homo sapiens chromosome 4 clone RP11-402c9 map 4, WORKING DRAFT SEQUENCE,
DE 25 unordered pieces.
XX
XX HTG; HTGS_DRAFT; HTGS_PHASE1.
XX
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Homidae; Homo.
XX
XX [1]
RN 1-193488
RA Birren B., Linton L., Nusbaum C., Lander E.,
RT "Homo sapiens chromosome 4, clone RP11-402c9",
RL Unpublished.
```


Db 127538 ACATATATTTGTGTATATACACATATATATGTGTATA 127498

RESULT 14
AX016326 1324 bp DNA linear PAT 07-SEP-2000
LOCUS
DEFINITION Sequence 29 from Patent WO9949046.
ACCESSION AX016326
VERSION AX016326.1 GI:10041881
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1324)
REFERENCE
AUTHORS Roberts,J.A., Wyatt,P. and Whitelaw,C.
TITLE Signal transduction protein involved in plant dehiscence
JOURNAL Patent: WO 9949046-A 29 30-SEP-1999;
ROBERTS JEREMY ALAN (GB); BIOGEMMA UK LTD (GB); WYATT PAUL (GB);
WHITELAW CATHERINE (GB)
FEATURES
SOURCE location/Qualifiers
1..1324
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
1200..>1322
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07687.1"
/db_xref="GI:10041882"
/translation="MATKSTGCTGTEKTSIEVKKLIINLYIDDDPLNRLHECHQ"
BASE COUNT 460 a 214 c 182 g 467 t 1 others
ORIGIN

Query Match 8.1%; Score 48.8; DB 6; Length 1324;
Best Local Similarity 71.3%; Pred. No. 0.11; Mismatches 17; Indels 18; Gaps 1;
Matches 87; Conservative 0;

QY 12 AATGAGATGGCAACAAATCCATGGAGATTCGAGAAATTAAGAAACTA----- 67
Db 1192 AATGAGATGGCAACAAATCCATGGAGATTCGAGAAATTAAGAAACTA----- 67
QY 68 -----AACGTGTGATCGTCGATGATGTCACATAACCTTATATTCGANG 113
Db 1252 AGAGAAACTATATCAAGCTGTGATGTCGATGATATATATCAATTAACCTTACACTCCACG 1311
QY 114 AG 115
Db 1312 AG 1313

RESULT 15
AC099331 151498 bp DNA linear PRI 12-FEB-2002
LOCUS
DEFINITION Homo sapiens chromosome 3 Clone RP11-90B15, complete sequence.
ACCESSION AC099331 AC079341
VERSION AC099331.2 GI:18652541
KEYWORDS
SOURCE HTG.
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 151498)
REFERENCE
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 151498)

AUTHORS
TITLE
JOURNAL
COMMENT

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
Direct Submission
Submitted (12-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Feb 12, 2002 this sequence version replaced gi:16674865.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
Drafting Center: BCM
----- Project Information
Center project name: chr-3
Center clone name: RP11-90B15 (bc0621)
----- Summary Statistics
Sequencing vector: unknown; 3% of reads
Sequencing vector: Plasmid; 68% of reads
Chemistry: Dye-terminator ET; 9% of reads
Chemistry: Dye-terminator Big Dye; 8% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 151408 bases at least Q40
Consensus quality: 151494 bases at least Q30
Consensus quality: 151498 bases at least Q20
Insert size: 149290; sum-of-contigs
Quality coverage: 9.0x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': Mapping in progress
3': RP11-613N24 (UWGC:bc0491) AC099357

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

	HindIII	BglII	EcoRI
SeqDerMap	FngPrint	SeqDerMap	FngPrint
-----	-----	-----	-----
970	964	8580	8813
-----	-----	-----	-----
6382	6463	2067	2050
-----	-----	-----	-----
512	<800	5474	5794
-----	-----	-----	-----
449	<800	6597	6842
-----	-----	-----	-----
			5809
-----	-----	-----	-----
			5961

7072	6880	793	783	2020	2012
7424	7336	13744	13846	3842	3873
13912	14004	1178	1161	490	<800
5543	6000	4238	4399	5591	5690
4549	4357	4137	4399	11077	10784
1890	1911	3381	3549	902	909
4597	4581	1527	1474	6995	6988
459	<800	1628	1604	6862	6988
2808	2860	748	783	435	<800
168	<800	1002	1006	4156	4128
1917	1911	810	783	11007	10784
1330	1291	5491	5794	7737	7651
3162	3222	74	<800	16863	17203
5149	5071	3542	3693	2501	2550
1889	1911	4519	4717	5219	5184
6123	6000	924	923	6649	6581
4466	4581	5553	5992	764	<800
3332	3222	1504	1474	2623	2550
3068	3222	1485	1474	1474	1446
2216	2301	1967	2050	527	<800
1112	1164	249	<800	1721	1688
4261	4208	990	1006	12425	12215
2258	2301	4000	4195	652	<800
5609	5624	763	783	437	<800
8713	8709	2195	2050	5996	5961
5532	5624	17584	17727	3302	3325
1030	1025	1952	1917	5889	5961
3181	3222	1006	1006	5653	5690
3337	3222	1462	1474	5122	5184
2523	2546	1881	1917	276	<800
492	<800	10522	10095		
7773	7793	7245	7567		
3705	3683	6941	7106		
1173	1164	248	<800		
4102	4065	2014	2050		
8385	8346	238	<800		

6792	6880	882	907
434	<800	1331	1316
		1753	1741
		2550	2406
		1166	1161
		1034	1006
		3982	4193
		7228	7567

FEATURES	
Source	Location/Qualifiers
	1. 151498
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-90B15"
	/clone_lib="RPC1 human BAC library 11"
misc_feature	40612..40696
	/note="Single subclone region"
misc_feature	119108..119321
	/note="Single subclone region"
BASE COUNT	42807 a 32520 c 31007 g 45164 t
ORIGIN	

Query Match 8.1%; Score 48.8; DB 9; Length 151498;
Best Local Similarity 51.9%; Pred. No. 0.09;
Matches 110; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY	379	GTAACCAAGGACGAATGATCCCTGTCATTAAACCAAGCATGSAATGCTGATGGAATATA	438
Db	114150	GTCATAGAAATTAATATTTCAACATATTCCTTTGGCATGTAAAGTCCATATGCCATATA	114209
QY	439	TATTTATTTATATGGAACACACATATATAAGCGCTAAGTGTATGTATGCATAGATACT	498
Db	114210	TATATGTGTGTGTGTATATATATATATACATATATATACGTATATATATACACTATATATA	114269
QY	499	TGCATGTGTGTGCTTTTACAAATTTAGGCTCTTATGTGTCGGATATATTAATCATGTAAG	558
Db	114270	CGGTGTGTGTGTGTGTATATATATATATATATACACAGCTAATATATATATATATA	114329
QY	559	TGTGTGCTTAAAGCTATATAAATATTTAAATA	590
Db	114330	TATATATATATATATATATATATATATATATATATA	114361

Search completed: November 7, 2002, 14:05:25
Job time : 2373 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 10:38:02 ; Search time 211 Seconds
(without alignments)
4922.906 Million cell updates/sec

Title: US-09-646-679-14
Perfect score: 605

Sequence: 1 ggcacgacgaatcgaaga.....aaataagggttcctctacc 605

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Database :

1	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1980.DAT *
N	Geneseq_0328021 *
2	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1981.DAT *
3	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1982.DAT *
4	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1983.DAT *
5	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1984.DAT *
6	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1985.DAT *
7	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1986.DAT *
8	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1987.DAT *
9	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1988.DAT *
10	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1989.DAT *
11	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1990.DAT *
12	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1991.DAT *
13	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1992.DAT *
14	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1993.DAT *
15	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1994.DAT *
16	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1995.DAT *
17	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1996.DAT *
18	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1997.DAT *
19	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1998.DAT *
20	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA1999.DAT *
21	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA2000.DAT *
22	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA2001.DAT *
23	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA2002.DAT *
24	/SIDS1/gcgdata1/geneseq/geneseqn-emb1/NA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	605	100.0	605	20	AAZ22974	Brassicica response
2	605	100.0	605	20	AAZ22977	Brassicica response
3	304.2	50.3	834	21	AAZC3077	Arabidopsis thaliana
4	297.4	49.2	818	20	AAZ22978	A. thaliana DZAT
5	296.8	49.1	700	21	AAZC47939	Arabidopsis thaliana
6	291.4	48.2	850	20	AAZC30139	Arabidopsis thaliana
7	289.4	47.8	576	20	AAZ22976	Brassicica response
8	106.2	17.6	1716	20	AAZ22975	Brassicica napus DZ
9	58.2	9.6	1656	21	AAZ79325	Pinus radiata cellu

C	10	48.8	8.1	1.324	20	AAZ22399	A. thaliana DZ2A73
C	11	45.6	7.5	1.965	20	AAK73166	Human immune/haema
C	12	44.5	7.5	9515	22	AA545453	Chemically pretreaa
C	13	44.8	7.4	32306	22	AAK89374	Human digestive sy
C	14	44.8	7.4	80595	20	AAV83939	HC-contig derived
C	15	44	7.3	4445	22	AAK76447	Human immune/haema
C	16	43.4	7.2	80240	20	AAK83940	NC-contig derived
C	17	43.4	7.2	106509	22	AA546359	Human immune/haema
C	18	43	7.1	7516	20	AA543060	Tumour suppressor
C	19	42.6	7.0	12661	22	AA513653	Human immune syste
C	20	42.6	7.0	26597	22	AA546748	Genomic DNA sequen
C	21	42.4	7.0	2304	21	AAZ43769	Tumour suppressor
C	22	42.4	7.0	8786	23	ABL16590	Human endometrialo
C	23	42	6.9	6541	24	ABL32315	Drosophila melanog
C	24	42	6.9	18855	24	ABL32315	Human immune syste
C	25	42	6.9	53075	22	AAK32610	Human immune syste
C	26	41.8	6.9	15418	21	AAK86671	Human immune/haema
C	27	41.6	6.9	15418	21	AA663785	Nucleotide sequenc
C	28	41.6	6.9	6244	24	ABL33484	Human immune syste
C	29	41.6	6.9	6244	24	ABL33485	Human immune syste
C	30	41.6	6.9	6971	24	ABL33237	Human immune syste
C	31	41.6	6.9	39068	22	AAK71820	Human immune syste
C	32	41.6	6.9	39068	22	AAK73078	Human immune/haema
C	33	41.6	6.9	39068	22	AAK86294	Human immune/haema
C	34	41.6	6.9	39110	22	AAK87844	Human immune/haema
C	35	41.6	6.9	39110	22	AAK71825	Human immune/haema
C	36	41.6	6.9	39110	22	AAK73087	Human immune/haema
C	37	41.4	6.8	6123	22	AAK87555	Human immune/haema
C	38	41.4	6.8	8952	22	AA543036	Human immune syste
C	39	41.4	6.8	13712	22	AA546445	Tumour suppressor
C	40	41.2	6.8	52562	22	ABL33531	Human immune syste
C	41	40.6	6.7	18060	24	AAK86659	Human immune/haema
C	42	40.2	6.6	2440	24	AAAD2313	Chemically treated
C	43	40.2	6.6	3540	24	AAZ45411	Pea DRR06 protein
C	44	40.2	6.6	11976	24	ABL32234	Human immune syste
C	45	40.2	6.6	18855	24	ABL33594	Human immune syste
C	46	40.2	6.6	18855	24	ABL33611	Human immune syste

ALIGNMENTS

RESULT	1
ID	AAZ22974
	AAZ22974 standard; DNA; 605 BP.
AC	AAZ22974;
DT	10-JAN-2000 (first entry)
XX	
DE	Brassica response regulator protein D22 encoding DNA.
XX	
XX	Signal transduction protein; dehiscence: male sterile plant; D22 gene;
XX	shatter resistance; oilseed rape; response regulator protein; ss.
OS	Brassica napus.
XX	
PN	WO949046-A1.
XX	
PD	30-SEP-1999.
XX	
PF	22-MAR-1999; 99WO-GB00905.
XX	
PR	20-MAR-1998; 98GB-0006113.
XX	
PA	(BIOG-) BIOGENMA UK LTD.
XX	
PI	Wyatt P, Roberts JA, Whiteleaw C;
XX	
DR	WPI: 1999-580449/49.
XX	
PT	P-PSDB; AA142637.
XX	
	A nucleic acid encoding a signal transduction protein involved in plant
	dehiscence, useful for producing shatter resistant male sterile plants

Claim 5; Fig 1; 71pp; English.

CC The invention provides a nucleic acid encoding a signal transduction
CC protein involved in the process of dehiscence. The nucleic acids and
CC proteins are useful for regulating or controlling dehiscence of a pod or
CC an anther in a plant, useful in the production of male sterile plants.
CC The methods, etc. may be used in production of shatter resistance or
CC shatter-delayed plants such as oilseed rape (*Brassica napus*). The present
CC sequence represents the DNA encoding a B. napus response regulator
CC protein DZ2.

Sequence 605 BP; 198 A; 101 C; 140 G; 166 T; 0 other;

Query Match	100.0%	Score 605;	DB 20;	Length 605;
Best Local Similarity	100.0%	Pred. No. 3.9e-159;		
Matches 605; Conservative	0;	Mismatches	0;	Indels 0; Caps 0

QY	1	GGACGGACGAGATGGAGATGGACACAAAATCCATGGAGATATCGAGAAATATTAAGAA	60
Db	1	GGCACGGACGAGATGGAGATGGACACAAAATCCATGGAGATATCGAGAAATATTAAGAA	60
QY	61	GAACTAAACGTGTGATCGTCGATGATGATCCACTAAACCTTATATTCATGAGAAAT	120
Db	61	GAACTAAACGTGTGATCGTCGATGATGATCCACTAAACCTTATATTCATGAGAAAT	120
QY	121	CATCAAAAGCATTTGGGGGTATTTACACAGACAGCAATTAACGGTGAAGAGGACATATCAT	180
Db	121	CATCAAAAGCATTTGGGGGTATTTACACAGACAGCAATTAACGGTGAAGAGGACATATCAT	180
QY	181	CCACCGTGACGGCGGCTATCTCTTTGACCTTATCTTATGATTAAGATAAAGAAATGCCGAGAG	240
Db	181	CCACCGTGACGGCGGCTATCTCTTTGACCTTATCTTATGATTAAGATAAAGAAATGCCGAGAG	240
QY	241	GGATGGTGTTTGACACACTAAGAAAGCTAAGAGAAATGGAAGTGAAGTCAATGATTTGTTGG	300
Db	241	GGATGGTGTTTGACACACTAAGAAAGCTAAGAGAAATGGAAGTGAAGTCAATGATTTGTTGG	300
QY	301	GGAGACTTCACGTGGCTGACCAATGAAGAGAGAGCCAGGGCTTTCATGAGAGCTGGACCTTAA	360
Db	301	GGAGACTTCACGTGGCTGACCAATGAAGAGAGAGCCAGGGCTTTCATGAGAGCTGGACCTTAA	360
QY	361	CCATTCCTTGGCAAAAACCGTTAACCAAGCAAGATCATCCCTCATTAACCAACTCAT	420
Db	361	CCATTCCTTGGCAAAAACCGTTAACCAAGCAAGATCATCCCTCATTAACCAACTCAT	420
QY	421	GGATGCTGATGATATATATTTATTTATTTATGGAACACACATTAATTAACGCTTAAGTGTG	480
Db	421	GGATGCTGATGATATATATTTATTTATTTATGGAACACACATTAATTAACGCTTAAGTGTG	480
QY	481	TATG	540
Db	481	TATG	540
QY	541	GATATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600
Db	541	GATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600
QY	601	CTACG 605	
Db	601	CTACG 605	
RESULT 2			
AAZ22977			
ID	AAZ22977 standard; DNA; 605 BP.		
XX	AAZ22977;		
XX	10-JAN-2000 (first entry)		
XX	Brassica response regulator protein D22 DNA sequence.		

XX Signal transduction protein; dehiscence; male sterile plant; Dv2 gene;
KW shatter resistance; oilseed rape; response regulator protein; ss.
KW

OS Brassica napus.

PN W09949046-A1.

PD 30-SEP-1999

PF 22-MAR-1999; 99WO-GB00905.

PR 20-MAR-1998; 98GB-0006113.

PA (BIOG-) BIOGEMMA UK LTD.

PI Wyatt P, Roberts JA, Whitelaw C;
.....

WPI; 1999-580449/49

XX

PT dehiscence, useful for producing shatter resistant male sterile plants

PS Example 2; Fig 6; 71pp; English.

CC The invention provides a nucleic acid encoding a signal transduction
CC protein involved in the process of dehiscentse. The nucleic acids and
CC proteins are useful for regulating or controlling dehiscentse of a pod or
CC an anther, in a plant, useful in the production of male sterile plants.
CC The methods, etc., may be used in production of shatter resistance or
CC shatter-delayed plants such as oilseed rape (*Brassica napus*). The present
CC sequence represents the B. napus D22 sequence.

50 Sequence 605 BP; 198 A; 101 C; 140 G; 166 T; 0 other;

Query Match	100.0%	Score 605;	DB 20;	Length 605;
Best Local Similarity	100.0%	Pred. No. 3.9e-159;		
Matches 605; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	GGCACGACGACGATGGAAGATGGCACACAAAATCCATGGGAGATATCGAGAAAATTAAGAA	60
Dp	1	GGGACGACGAGGATGGAAGATGGCACACAAAATCCATGGGAGATATCGAGAAAATTAAGAA	60
QY	61	GAAACTAAACGGTGTGATCGTCGATGATGATCACTAAACCTTATATTCATGGAAGAT	120
Dp	61	GAAACTAAACGGTGTGATCGTCGATGATGATCACTAAACCTTATATTCATGGAAGAT	120
QY	121	CATCAAAAGCGATTGGGGGTATTTCACAGACAGCAATTAAGGTGAGGCGATATCAT	180
Dp	121	CATCAAAAGCGATTGGGGGTATTTCACAGACAGCAATTAAGGTGAGGCGAGATATCAT	180
QY	181	CCACGGTGAAGGCGGCTCATCTTTTGAACCTATATCCATTAAGATTAAGAAATGCCCGAG	240
Dp	181	CCACGGTGAAGGCGGCTCATCTTTTGAACCTATATCCATTAAGATTAAGAAATGCCCGAG	240
QY	241	GGATGGTGTGTTGGACAACTAAGAAAGCTAAGAGAAATGGAAGTAAGTCAATGATTGTGG	300
Dp	241	GGATGGTGTGTTGGACAACTAAGAAAGCTAAGAGAAATGGAAGTAAGTCAATGATTGTGG	300
QY	301	GGGATGTTACGTGGTGTGACAACTAAGAGGAGGCCAAGGCTTTCATGGAAAGTGGACCTAA	360
Dp	301	GGGATGTTACGTGGTGTGACAACTAAGAGGAGGCCAAGGCTTTCATGGAAAGTGGACCTAA	360
QY	361	CCATTGCTGGGCAAAACCGTTAACCAAGGACAGATCACTCCCTCATTAACCAACATCAT	420
Dp	361	CCATTGCTGGGCAAAACCGTTAACCAAGGACAGATCACTCCCTCATTAACCAACATCAT	420
QY	421	GGATGCTTATGATATATATTTATATATATATGAAAACACATTAATTAACGCTTAAGTGTG	480
Dp	421	GGATGCTTATGATATATATTTATATATATATGAAAACACATTAATTAACGCTTAAGTGTG	480

[illegible]

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142927.
PR 14-JUL-1999; 99US-0143544.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160757.
PR 21-OCT-1999; 99US-0160758.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161982.
PR 29-OCT-1999; 99US-0162122.

Query Match 49.1%; Score 296.8; DB 21; Length 700;
Best Local Similarity 81.9%; Pred. No. 4.6e-73;
Matches 362; Conservative 0; Mismatches 62; Indels 18; Gaps 1;

QY 12 AATGAAAGATGGCAAAATCCATGGAGATATCCGAGAAATTAAGAAAGAACTA----- 67
DB 88 AATCGAATGGCAAAATCCCGAGAGTACCGAGAAACCAAGTGAAGAGTGA 147
QY 68 -----AACGTGTGATCGTCGATGATGATCCACTAACTTAATATTCATG 113
DB 148 AGAAGAACTAATCAACGTGTGATGTCGATGATGATCATTAACCTAGACTCAG 207
QY 114 AGAAGATCATCAAGCCATTTGGGGTATTTCAAGACAGAGAAATACGTTGAGAGCGAG 173
DB 208 AGATGATCATCAAAACGATGGAGGATTTCTGACTCAAGAGATGGCGAGAGCGAG 267
QY 174 TAATCATCAACCGTGAAGCGGCTCACTTTGACCTTAATCCATGATGATGAAGAAATGC 233
DB 268 TGATCTCCACCGTGAAGCGGAGAGCATCTTCCACTTATCTAAAGGATAAGAAATGC 327
QY 234 CCGAGAGGATGTGTTTGCACAACTAAGAACTAAGAAATGAAGTGAAGTGAATGA 293

```
Db      328  CTGAGAGGAGTGGAGTTTCACACACTAGACCTAGAGAAATGAAGTGCATCATGA 387
OY      294  TTGTTGGGGGACTCTACTGCTGCACATGAGAGAGGCGGCTTCATGSAAGCTG 353
Db      388  TCGTTGGGGTACCTGCTGCTGACCTGACAGAGAGAGCGCTTATGGAAGCTG 447
OY      354  GACTTAACCATTTGCTTGGCAAAACCGTTAACAGAGACAAGATCATCCCTCATTAACC 413
Db      448  GGCTCAACCATTTGCTTGGAAAAACCGTTAACAGAGCAAGATCTCCCGCTCATTAAGCC 507
OY      414  AACCTACGATGCTGATGAT 435
Db      508  ACCTCTCGATGCTTGAATGAT 529

RESULT 6
AAC39139
XX      ID    AAC39139 standard; DNA: 850 BP.
AC      AAC39139;
DF      17-OCT-2000 (first entry)
XX      DE    Arabidopsis thaliana DNA fragment SEQ ID NO: 23516.
XX      KW    Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway;
KW      metabolic pathway; promoter; termination sequence; ss.
XX      OS    Arabidopsis thaliana.
PN      EP1033405-A2.
PD      06-SEP-2000.
XX      PE    25-FEB-2000; 2000EP-0301439.
XX      PR    25-FEB-1999; 99US-0121825.
XX      PR    05-MAR-1999; 99US-0123180.
XX      PR    09-MAR-1999; 99US-0123548.
XX      PR    23-MAR-1999; 99US-0125788.
XX      PR    25-MAR-1999; 99US-0126264.
XX      PR    29-MAR-1999; 99US-0126785.
XX      PR    01-APR-1999; 99US-0127462.
XX      PR    06-APR-1999; 99US-0128234.
XX      PR    08-APR-1999; 99US-0128714.
XX      PR    16-APR-1999; 99US-0129845.
XX      PR    19-APR-1999; 99US-0130077.
XX      PR    21-APR-1999; 99US-0130449.
XX      PR    23-APR-1999; 99US-0130510.
XX      PR    28-APR-1999; 99US-0130891.
XX      PR    30-APR-1999; 99US-0131449.
XX      PR    30-APR-1999; 99US-0132048.
XX      PR    04-MAY-1999; 99US-0132407.
XX      PR    04-MAY-1999; 99US-0132484.
XX      PR    05-MAY-1999; 99US-0132485.
XX      PR    06-MAY-1999; 99US-0132486.
XX      PR    07-MAY-1999; 99US-0132487.
XX      PR    11-MAY-1999; 99US-0134256.
XX      PR    14-MAY-1999; 99US-0134218.
XX      PR    14-MAY-1999; 99US-0134219.
XX      PR    14-MAY-1999; 99US-0134221.
XX      PR    14-MAY-1999; 99US-0134370.
XX      PR    18-MAY-1999; 99US-0134768.
XX      PR    19-MAY-1999; 99US-0134941.
XX      PR    20-MAY-1999; 99US-0135124.
XX      PR    21-MAY-1999; 99US-0135353.
XX      PR    24-MAY-1999; 99US-0135639.
XX      PR    25-MAY-1999; 99US-0136021.
XX      PR    27-MAY-1999; 99US-0136392.
XX      PR    28-MAY-1999; 99US-0136782.

PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140891.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0143297.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145313.
PR      27-JUL-1999; 99US-0145318.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
```



```

XX      Wyatt P, Roberts JA, Whitelew C;
PI
XX
XX      WPI: 1999-580449/49.
DR      P-PSDB; AA42652.
XX
XX      A nucleic acid encoding a signal transduction protein involved in plant
PT      dehiscence, useful for producing shatter resistant male sterile plants
PT
XX
XX      Example 2; Fig 6; 71pp; English.
PS
XX
XX      The invention provides a nucleic acid encoding a signal transduction
CC      protein involved in the process of dehiscence. The nucleic acids and
CC      proteins are useful for regulating or controlling dehiscence of a pod or
CC      an anther in a plant, useful in the production of male sterile plants.
CC      The methods, etc. may be used in production of shatter resistance or
CC      shatter delayed plants such as oilseed rape (Brassica napus). The present
CC      sequence represents the B. napus D22B cDNA sequence.
XX
XX      Sequence 576 BP; 188 A; 102 C; 122 G; 161 T; 3 other;
SQ
XX
XX      Query Match          47.8%; Score 289.4; DB 20; Length 576;
XX      Best Local Similarity 75.4%; Pred. No. 5e-71;
XX      Matches 416; Conservative 0; Mismatches 94; Indels 42; Gaps 3;
XX
XX      78 TCGTCGATGATGATCCACTTAACCTTATATTCATGAGAGATCATCAAGCGATTGGGG 137
DB      1 TCGTCNATGATGATCCTGTAATACGTAACCTTACAGAGATATCATCAATCAACGCGTG 60
XX
XX      138 GTATTTCCACAGACGCGAATTAAGGTGAGGAGGAGGAGGAGTATCCACCGCTGCGGCGCT 197
DB      61 GAATTTCCACAGACGCGAATTAAGGTGAGGAGGAGGAGGAGTATCCACCGCGCGCAATG 120
XX
XX      198 CATCTTTGACCTTATCTTAATGATTAAGAAATGCCGAGAGGAGGTGTTTCGACAA 257
DB      121 CATCTTGCACCTTATCTTAATGATTAAGAAATGCCGAGAGGAGGTGACTTTCGCGCA 180
XX
XX      258 CTAAAGAGCTTAAGAAATGGAAGTGAATGATGTTGGGGTCACTTCAGTGGCTG 317
DB      181 CTAAAGAGCTTAAGAAATGGAAGTGAATGATGTTGGGGTCACTTCAGTGGCTG 240
XX
XX      318 ACAATGAAGAGAGCGGAGGCTTTCATGAGAGTGAATTAACCATGCTTGGCAAAAC 377
DB      241 ACAATGAAGAGAGCGGAGGCTTTCATGAGAGTGAATTAACCATGCTTGGCAAAAC 300
XX
XX      378 CGTTAAOCAGAGCAGATCACCCTTCATTAACCAACTCATGATGATGATGGA--- 434
DB      301 CGTTAAOCAGAGCAGATCACCCTTCATTAACCAACTCATGATGATGATGGA 360
XX
XX      435 -----TATATTTTATATATGGAACACACATTAATACGT 471
DB      361 ATGAATTCGCGCACTACATATCTACATTAATTAAGAAACACATTAATTAACGT 420
XX
XX      472 -----CTAAGTGTATATGATGATGATGATGATGATGATGATGATGATGATG 524
DB      421 CATACACCTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 480
XX
XX      525 GTT-----CTTATGTCGCGGATATATATATATATATATATATATATATATAT 575
DB      481 GTTGTATTTTATATGTCGCGGATATATATATATATATATATATATATATATAT 540
XX
XX      576 TAAATATATTTAA 587
DB      541 TAAATATATGAA 552

```

```

DT      10-JAN-2000 (first entry)
XX
XX      Brassica napus D22B promoter region and partial coding sequence.
DE
XX
XX      Signal transduction protein; dehiscence; male sterile plant; D22B gene;
KW      shatter resistance; oilseed rape; response regulator protein; ss.
XX
XX      Brassica napus.
OS
XX
XX      MO949046-A1.
PN
XX
XX      30-SEP-1995.
PD
XX
XX      22-MAR-1999; 99WO-GB00905.
PF
XX
XX      20-MAR-1998; 98GB-0006113.
PR
XX
XX      (BIOG-) BIOGEMMA UK LTD.
PA
XX
XX      Wyatt P, Roberts JA, Whitelew C;
PI
XX
XX      WPI: 1999-580449/49.
DR      P-PSDB; AA42644.
XX
XX      A nucleic acid encoding a signal transduction protein involved in plant
PT      dehiscence, useful for producing shatter resistant male sterile plants
PT
XX
XX      Example 2; Fig 5; 71pp; English.
PS
XX
XX      The invention provides a nucleic acid encoding a signal transduction
CC      protein involved in the process of dehiscence. The nucleic acids and
CC      proteins are useful for regulating or controlling dehiscence of a pod or
CC      an anther in a plant, useful in the production of male sterile plants.
CC      The methods, etc. may be used in production of shatter resistance or
CC      shatter delayed plants such as oilseed rape (Brassica napus). The present
CC      sequence represents the promoter region and a partial coding region of
CC      B. napus D22B gene.
XX
XX      Sequence 1716 BP; 604 A; 256 C; 264 G; 576 T; 16 other;
SQ
XX
XX      Query Match          17.6%; Score 106.2; DB 20; Length 1716;
XX      Best Local Similarity 83.9%; Pred. No. 1.1e-19;
XX      Matches 120; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
XX
XX      51 AATTAAGAGAGAACTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 110
DB      1568 AAGTGAAGAGAGAACTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1627
XX
XX      111 ATGAGAGATCATCAAAAGGATTTGGGGATTTTCACAGACAGCAATTAACGATGAGGAG 170
DB      1628 ACGAGAGATCATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1687
XX
XX      171 CAGTATCATTCACCGCTGACGCG 193
DB      1688 CAGTGAACATCCACGCGACGCG 1710

```

RESULT 8
 AA222975
 ID AA222975 standard; DNA; 1716 BP.
 XX
 AC AA222975;
 XX

RESULT 9
 AAA79325
 ID AAA79325 standard; cDNA; 1656 BP.
 XX
 AC AAA79325;
 XX
 DT 27-NOV-2000 (first entry)
 XX
 DE Pinus radiata cell signalling involved polynucleotide seq ID NO:63.
 XX
 KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism; ss.
 XX

```

XX XX Signal transduction protein; dehiscence; male sterile plant; DZAT3;
KW Shatter resistance; oilseed rape; response regulator protein;
KW promoter; ss.
OS Arabidopsis thaliana.
PN MO9949046-A1.
XX
XX 30-SEP-1999.
XX
XX 22-MAR-1999; 99MO-GB00905.
XX
XX PR 20-MAR-1998; 98GB-0006113.
XX
XX PA (BIOG-) BIOGENMA UK LTD.
XX
XX PI Wyatt P, Roberts JA, Whiteleaw C;
XX
XX DR WP1: 1999-580449/49.
XX
XX P-PSDB: AAY42648.
XX
XX A nucleic acid encoding a signal transduction protein involved in plant
PT dehiscence, useful for producing shatter resistant male sterile plants
PT
XX
XX Example 3; Fig 11; 71pp; English.
XX
XX The invention provides a nucleic acid encoding a signal transduction
CC protein involved in the process of dehiscence. The nucleic acids and
CC proteins are useful for regulating or controlling dehiscence of a pod or
CC an anther in a plant, useful in the production of male sterile plants.
CC The methods, etc. may be used in production of shatter resistance or
CC shatter delayed plants such as oilseed rape (Brassica napus). The present
CC sequence represents the promoter region and a partial coding region of
CC A. thaliana DZAT3 gene.
XX
XX
SQ Sequence 1324 BP; 460 A; 214 C; 182 G; 467 T; 1 other;

Query Match      8.1%; Score 48.8; DB 20; Length 1324;
Best Local Similarity 71.3%; Pred. No. 0.001;
Matches 87; Conservative 0; Mismatches 17; Indels 18; Gaps 1;

QY 12 AATCGAAGATGCGACAAATCCATGCGAGATATCGAGAAATTAAGAAACTA---- 67
    |||
Db 1192 AATCGAAGATGCGACAAATCCATGCGAGATATCGAGAAATTAAGAAACTAAGTGA 1251
    |||

QY 68 -----AACGTTGATGTCGATGATGATCCACTAAACCTTAATTCATG 113
    |||
Db 1252 AGAAGAACTAATCAACGTTGATGATGATGATCCACTTAACCTTAAGTCCACAG 1311
    ||
QY 114 AG 115
Db 1312 AG 1313

RESULT 11
AAK73166/c
ID ID AAK73166 standard; DNA; 19965 BP.
XX
XX AAK73166.
XX
XX AC
XX DT
XX 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27978.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX MO200157182-A2.
XX
XX PN
XX

```


PT	Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
PR	useful for preventing, diagnosing and/or treating cancers and
PT	metastasis -
XX	
XX	
XX	Disclosure; SEQ ID NO 27978; 3071bp + Sequence Listing; English.
XX	
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC	amino acid sequences given in AAK62170 to AAK1921. (I) have cytostatic
CC	activity and can be used in gene therapy and vaccine production. (I)
CC	proteins and polynucleotides may be used in the prevention, diagnosis and
CC	treatment of diseases associated with inappropriate (I) expression. For
CC	example, they may be used to treat disorders associated with decreased
CC	expression by rectifying mutations or deletions in a patient's genome
CC	that affect the activity of (I) by expressing inactive proteins or to
CC	supplement the patient's own production of (I). Additionally, (I)
CC	polynucleotides may be used to produce the secreted (I), by inserting the
CC	nucleic acids into a host cell and culturing the cell to express the
CC	protein. (I) proteins and polynucleotides may be used to prevent,
CC	diagnose and treat immune/haematopoietic-related diseases, especially
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC	to AAK87694 represent human immune/haematopoietic antigen genomic
CC	sequences from the present invention. AAK51942 to AAK54950 and AAK82169
CC	represent sequences used in the exemplification of the present invention.
XX	
SQ	Sequence 19965 BP; 4699 A; 5296 C; 5505 G; 4465 T; 0 other;
Query Match	7.5%; Score 45.6; DB 22; Length 19965;
Best Local Similarity	54.9%; Pred. No. 0.022;
Matches	90; Conservative 0; Mismatches 74; Indels 0; Gaps 0
OY	427 TTGATGGATATATATTTATATTATTTGAAACACACATATATACGTCTAAGTGTATGTA 486
Db	4721 TTCTATATATATATTTTTCATATATATATATATATATATATATATATATTTATATAT 4662
OY	487 TGCATATACCTGCAGCTGCTGCTTTTGAATTTAGGCTCTTATACGCCGATATA 546
Db	4661 TTCTATATATATTTTATATATATATTTTCATATATATATTTTTCATATATATATATATA 4502
OY	547 TAAATCATGTAAGTGTGCTTTAGCTTATTAATAATTTAATA 590
Db	4601 TATTCATATATATATTTATATATATTTTTCATATATATATTTTTCATATATATATTTTAA 4558
RESULT 12	
AAK545453	
ID	AAK545453 standard; DNA: 9515 BP.
XX	
AC	AAK545453;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Chemically pretreated complementary DNA associated with cell cycle #79.
XX	
KW	Cell cycle; human; Cdc dinucleotide; cytosine methylation; HIV; aging;
KW	human immunodeficiency virus; neurodegenerative disorder; solid tumour;
KW	graft-versus-host disease; glomerular disease; lewy body disease; cancer;
KW	arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
KW	immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
XX	PCR primer.
XX	
OS	Homo sapiens.
XX	
PN	WO2001.68911-A2.
XX	
PD	20-SEP-2001.
XX	
PF	15-MAR-2001; 2001WO-EP02945.
XX	
PR	15-MAR-2000; 2000DE-1013847.
PR	06-APR-2000; 2000DE-1019058.
PR	07-APR-2000; 2000DE-1019173.
PR	30-JUN-2000; 2000DE-1032529.

PB			01-SEP-2000; 2000DE-1043826.
XX			(EPIC-) EPIGENOMICS AG.
PA	Olek A,	Piepenbrock C,	Berlin K;
PI	WPI; 2001-602751/68.		
DR			
XX			
PT	Designing primers and probes for analysing diseases associated with cytosine methylation state e.g. arthritis, cancer, aging, arteriosclerosis comprising fragments of chemically modified genes associated with cell cycle -		
PS	Claim 1; SEQ ID No 159; 28pp; English.		
CC			
CC			
CC	Sequences AA645296-AA645520 represent chemically pretreated genomic DNA molecules associated with the cell cycle and specific PCR primers of the invention. The sequences are useful for detecting the methylation state of all CpG dinucleotides in a sequence and therefore for analysing associated diseases. By analysing cytosine methylations in the pretreated DNA, genetic and/or epigenetic parameters for the diagnosis and therapy of existing diseases or the predisposition to specific diseases can be ascertained. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The sequences of the invention are useful for the diagnosis and therapy of HIV infection, neurodegenerative disorders, graft-versus-host disease, aging, glomerular disease, lewy body disease, arthritis, arteriosclerosis, solid tumours and cancers.		
CQ			
SQ	Sequence 9515 BP; 2719 A; 220 C; 1904 G; 4672 T; 0 other:		
	Query Match 7.5%; Score 45.4; DB 22; Length 9515; Best Local Similarity 50.7%; Pred. No. 0.019;		
	Matches 109; Conservative 0; Mismatches 106; Indels 0; Gaps 0		
OY	375 AACCGTTAACAGACAGATCATGCCCTATTACCACACTCATGTGGTGAATGA 434 DB 4665 AAGTTTTAGAAGACAAGAAAAAATTAATTCGTATAATAGAGTTTGAAATTTATATGA 4724 OY 435 TAATATTTATATTTATANGCAACACATATAACGTCFAAGTCGTATGATGCATAGA 494 DB 4725 TATTTATATTTAAAATTATTAANTCGRATTTATATTATTTATATTAATAGTATA 4784 OY 495 TACTTGATGTGTGTGTTTAGAATTAGGCTCTTATCGTCGTATATATATCATG 554 DB 4785 TAGATATTTTGGCTTATATAGTTGTATAGATTTAAAGTGNATTATTAAGAAGTTGTT 4844 OY 555 TAAGTTGTGCTTTAAGCTTATAAATTTTAAT 589 DB 4845 TAGGGTATTTAATATATATATATTTATTTAATT 4879 		
RESULT 13			
ID	AAK89374/C		
XX	AAR89374 standard; DNH; 32206 BP.		
AC	AAK89374;		
XX			
DT	05-NOV-2001 (first entry)		
DE	Human digestive system antigen genomic sequence SHQ ID NO: 2950.		
KX	Human: digestive system antigen; gene therapy; cancer; appendicitis;		
KM	ulcerative colitis; Infection; Hirschsprung's disease; chronic colitis;		
RW	digestive system disorder; Meckel's diverticulum; ds.		
XX			
OS	Homo sapiens.		
PM	WO200155314-A2.		
XX			
PD	02-AUG-2001.		
XX			

PE	17-JAN-2001.	2001WO-US01324.
XX		
PR	31-JAN-2000.	2000US-0179065.
PR	04-FEB-2000.	2000US-0180628.
PR	24-FEB-2000.	2000US-0184664.
PR	02-MAR-2000.	2000US-0186350.
PR	16-MAR-2000.	2000US-0189874.
PR	17-MAR-2000.	2000US-0190076.
PR	18-MAR-2000.	2000US-0198122.
PR	19-MAR-2000.	2000US-0205515.
PR	07-JUN-2000.	2000US-0209467.
PR	28-JUN-2000.	2000US-0214886.
PR	30-JUN-2000.	2000US-0215135.
PR	07-JUL-2000.	2000US-0216647.
PR	07-JUL-2000.	2000US-0216880.
PR	11-JUL-2000.	2000US-0217487.
PR	11-JUL-2000.	2000US-0217496.
PR	14-JUL-2000.	2000US-0218290.
PR	26-JUL-2000.	2000US-0220963.
PR	26-JUL-2000.	2000US-0220964.
PR	14-AUG-2000.	2000US-0224518.
PR	14-AUG-2000.	2000US-0224519.
PR	14-AUG-2000.	2000US-0225213.
PR	14-AUG-2000.	2000US-0225214.
PR	14-AUG-2000.	2000US-0225266.
PR	14-AUG-2000.	2000US-0225267.
PR	14-AUG-2000.	2000US-0225268.
PR	14-AUG-2000.	2000US-0225270.
PR	14-AUG-2000.	2000US-0225447.
PR	14-AUG-2000.	2000US-0225757.
PR	14-AUG-2000.	2000US-0225758.
PR	14-AUG-2000.	2000US-0225759.
PR	18-AUG-2000.	2000US-0226279.
PR	22-AUG-2000.	2000US-0226681.
PR	22-AUG-2000.	2000US-0226686.
PR	22-AUG-2000.	2000US-0227182.
PR	23-AUG-2000.	2000US-0227009.
PR	30-AUG-2000.	2000US-0228924.
PR	01-SEP-2000.	2000US-0229287.
PR	01-SEP-2000.	2000US-0229343.
PR	01-SEP-2000.	2000US-0229344.
PR	01-SEP-2000.	2000US-0229345.
PR	05-SEP-2000.	2000US-0229509.
PR	05-SEP-2000.	2000US-0229513.
PR	06-SEP-2000.	2000US-0230437.
PR	06-SEP-2000.	2000US-0230438.
PR	08-SEP-2000.	2000US-0231242.
PR	08-SEP-2000.	2000US-0231243.
PR	08-SEP-2000.	2000US-0231244.
PR	08-SEP-2000.	2000US-0231413.
PR	08-SEP-2000.	2000US-0231414.
PR	08-SEP-2000.	2000US-0232080.
PR	08-SEP-2000.	2000US-0232081.
PR	12-SEP-2000.	2000US-0231968.
PR	14-SEP-2000.	2000US-0233397.
PR	14-SEP-2000.	2000US-0233398.
PR	14-SEP-2000.	2000US-0233399.
PR	14-SEP-2000.	2000US-0234200.
PR	14-SEP-2000.	2000US-0234201.
PR	14-SEP-2000.	2000US-0233063.
PR	14-SEP-2000.	2000US-0233064.
PR	14-SEP-2000.	2000US-0233065.
PR	21-SEP-2000.	2000US-0234223.
PR	21-SEP-2000.	2000US-0234274.
PR	25-SEP-2000.	2000US-0234997.
PR	25-SEP-2000.	2000US-0234998.
PR	26-SEP-2000.	2000US-0235484.
PR	27-SEP-2000.	2000US-0235834.
PR	27-SEP-2000.	2000US-0235836.
PR	29-SEP-2000.	2000US-0236327.
PR	29-SEP-2000.	2000US-0236367.
PR	29-SEP-2000.	2000US-0236368.
PR	29-SEP-2000.	2000US-0236369.
PT		
PR	29-SEP-2000.	2000US-0236370.
PR	02-OCT-2000.	2000US-0236802.
PR	02-OCT-2000.	2000US-0237037.
PR	02-OCT-2000.	2000US-0237038.
PR	02-OCT-2000.	2000US-0237039.
PR	02-OCT-2000.	2000US-0237040.
PR	13-OCT-2000.	2000US-0239935.
PR	13-OCT-2000.	2000US-0239937.
PR	20-OCT-2000.	2000US-0240960.
PR	20-OCT-2000.	2000US-0241221.
PR	20-OCT-2000.	2000US-0241785.
PR	20-OCT-2000.	2000US-0241786.
PR	20-OCT-2000.	2000US-0241787.
PR	20-OCT-2000.	2000US-0241808.
PR	20-OCT-2000.	2000US-0241809.
PR	20-OCT-2000.	2000US-0241826.
PR	01-NOV-2000.	2000US-0244617.
PR	08-NOV-2000.	2000US-0246474.
PR	08-NOV-2000.	2000US-0246475.
PR	08-NOV-2000.	2000US-0246476.
PR	08-NOV-2000.	2000US-0246477.
PR	08-NOV-2000.	2000US-0246478.
PR	08-NOV-2000.	2000US-0246523.
PR	08-NOV-2000.	2000US-0246524.
PR	08-NOV-2	

PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases .
 PS Disclosure: SEQ ID NO 2950; 986bp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a genomic DNA fragment
 CC encoding a digestive system antigen of the invention.
 XX
 SQ Sequence 32206 BP; 9503 A; 5480 C; 7017 G; 10206 T; 0 other;
 Query Match 7.4%; Score 44.8; DB 22; Length 32206;
 Best Local Similarity 54.2%; Pred. No. 0.043;
 Matches 91; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
 QY 430 ATGATATATATTTTATTTATGGAACACACATATATAGCTTAAGTGTATGATGAC 489
 Db 22201 ATGATATATATACATGATATATATATGATGATATATACATGATGATATATATGAT 22142
 QY 490 ATAGATACCTGCATGCTGCTGTTTGAATTTAGGCTCTTATCGTCCGATATATAA 549
 Db 22141 ATATATACATGATGATATATGATGATATATATACATGATGATGATGATATATAT 22082
 QY 550 TCATGATGTTGCTTAAAGCTTAAATATTTAATAAGGTTT 597
 Db 22081 ACATGCTGATGATGATGATATATGCTGATATATATATATATATATGCTGT 22034
 RESULT 14
 AAV83939
 ID AAV83939 standard; DNA; 80595 BP.
 AC AAV83939;
 XX
 DT 03-MAR-1999 (first entry)
 XX
 DE HC-contig derived from normal human chromosome 10q25.2 region.
 XX
 KW Yeast artificial chromosome; YAC; probe: eukaryotic chromosome;
 KW neocentromere; replication; extra-chromosomal element; segregation;
 KW cell division; artificial chromosome; gene therapy; mardel1(10);
 KW human artificial chromosome; transgenic; chromosome 10; 10q25.2; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9851790-A1.
 XX
 PD 19-NOV-1998.
 XX
 PF 13-MAY-1998; 98WO-AU00352.
 XX
 PR 26-AUG-1997; 97AU-0008791.
 PR 13-MAY-1997; 97AU-0006784.
 XX
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 XX
 PI Cancilla MR, Choo K, Du Sart D;
 XX
 DR WPI: 1999-009773/01.
 XX
 PT New isolated nucleic acid comprising neocentromere sequences from
 PT eukaryotic chromosome - used to produce replicable, segregating
 PT artificial chromosomes that can carry large amounts of DNA for gene
 PT therapy
 XX
 PS Claim 8; Fig 6; 540pp; English.
 CC The present sequence represents the HC-contig derived from normal human
 CC chromosome 10, 10q25.2 region. This region can be naturally mutated to

CC produce an unusual chromosomal marker referred to as mardel(10). The
 CC mardel(10) marker is mitotically stable and contains a functional
 CC neocentromere at a location regarded as non-centromeric. This
 CC neocentromere maps to q25.2 on chromosome 10. The specification describes
 CC nucleic acid sequences derived from a eukaryotic chromosome, including a
 CC neocentromere or its functional derivative or hybrid, that are able, in
 CC a competent cell, of replicating, acting as extra-chromosomal element
 CC and segregating during cell division. The sequences can be used to
 CC construct artificial chromosomes for use in gene therapy comprising a
 CC replicable, segregating nucleic acid that confers a specific phenotype
 CC on cells. Human artificial chromosomes can propagate in human cells and
 CC carry large amounts of DNA (e.g. therapeutic genes), and, being
 CC extra-chromosomal, they are not mutagenic. The artificial chromosomes
 CC are also useful for generation of transgenic plants and animals, in
 CC production of proteins and to make diagnostic reagents, e.g. for
 CC expression of cytokines, receptors and growth factors, or to increase
 CC the copy number of a gene in a cell. The constructs may also be
 CC used for functional and structural analysis of chromosomes.
 XX
 SQ Sequence 80595 BP; 23183 A; 16613 C; 16824 G; 23975 T; 0 other;
 Query Match 7.4%; Score 44.8; DB 20; Length 80595;
 Best Local Similarity 54.2%; Pred. No. 0.06;
 Matches 91; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
 QY 430 ATGATATATATTTTATTTATGGAACACACATATATAGCTTAAGTGTATGATGAC 489
 Db 15588 ATACATTAATATGATATATATGATGATATATGATATATATGATATATATATAC 15647
 QY 490 ATAGATACCTGCATGCTGCTGTTTGAATTTAGGCTCTTATCGTCCGATATATAA 549
 Db 15648 ATAAATATGATATATATGCTATATAGACATATAATATGATATATATATATAGACATA 15707
 QY 550 TCATGATGTTGCTTAAAGCTTAAATATTTAATAAGGTTT 597
 Db 15708 ATATGATATATATGCTTATATAGACATATAATATGATATATATATATATATAT 15755
 RESULT 15
 AAK76447
 ID AAK76447 standard; DNA; 4445 BP.
 AC AAK76447;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:31259.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0180076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 10:50:27 ; Search time 47 Seconds
(without alignments)
3161.878 Million cell updates/sec

Title: US-09-646-679-14

Perfect score: 605

Sequence: 1 ggcacgacgaatcgaaga.....aataaggttcctacc 605

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents NA: *
1: /cgn2_6/pdata/1/ina/5A.COMB.seq: *
2: /cgn2_6/pdata/1/ina/6A.COMB.seq: *
3: /cgn2_6/pdata/1/ina/6A.COMB.seq: *
4: /cgn2_6/pdata/1/ina/6B.COMB.seq: *
5: /cgn2_6/pdata/1/ina/6C.COMB.seq: *
6: /cgn2_6/pdata/1/ina/6D.COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	58.2	9.6	US-09-228-986-63	Sequence 63, Appl
2	34	8.9	US-08-232-463-14	Sequence 14, Appl
3	44.8	7.4	US-09-078-294-3	Sequence 3, Appl
4	44	80246	US-09-078-294-4	Sequence 4, Appl
5	40	5014	US-09-381-862-3	Sequence 3, Appl
6	39.8	6.6	US-08-486-013-70	Sequence 70, Appl
7	39.8	6.6	US-08-482-279-70	Sequence 70, Appl
8	39.8	6.6	US-08-342-268-70	Sequence 70, Appl
9	39.8	6.6	US-09-013-968-70	Sequence 70, Appl
10	39.8	6.6	US-09-397-386-70	Sequence 70, Appl
11	39.8	6.6	US-09-031-563-26	Sequence 26, Appl
12	38.6	6.4	US-09-078-294-4	Sequence 4, Appl
13	38.6	6.4	US-09-078-294-4	Sequence 4, Appl
14	37.6	6.2	US-08-629-643A-5	Sequence 5, Appl
15	37.6	6.2	US-08-280-799-1	Sequence 1, Appl
16	37.6	6.2	US-09-155-884-5	Sequence 5, Appl
17	35.4	5.9	5312912-3	Patent No. 5312912
18	34.6	5.7	US-08-484-105-7	Sequence 7, Appl
19	34.6	5.7	US-08-484-106-7	Sequence 7, Appl
20	34.6	5.7	US-08-136-743B-3	Sequence 3, Appl
21	34.2	5.7	US-08-771-784-1	Sequence 1, Appl
22	34.2	5.7	US-09-076-756-1	Sequence 1, Appl
23	34.2	5.7	US-09-255-984-1	Sequence 1, Appl
24	33.8	5.6	US-09-797-906-3	Sequence 3, Appl
25	33.6	5.6	US-08-936-135-9	Sequence 9, Appl
26	33.6	5.6	US-08-936-135-11	Sequence 11, Appl
27	33.6	5.6	US-08-936-135-13	Sequence 13, Appl

28	33.6	5.6	4784	3	US-08-936-135-15	Sequence 15, Appl
29	33.6	5.6	6152	4	US-08-973-462-1	Sequence 1, Appl
30	33.4	5.5	3098	1	US-08-447-500-1	Sequence 1, Appl
31	33.4	5.5	3098	1	US-08-454-097-1	Sequence 1, Appl
32	33.4	5.5	3098	1	US-08-447-408-1	Sequence 1, Appl
33	33.4	5.5	3098	1	US-08-453-866-1	Sequence 1, Appl
34	33.4	5.5	3098	3	US-08-185-359-1	Sequence 1, Appl
35	33.2	5.5	288	4	US-09-556-868-1	Sequence 1, Appl
36	33.2	5.5	2061	2	US-08-960-022-11	Sequence 11, Appl
37	32.6	5.4	870	2	US-08-420-629-9	Sequence 9, Appl
38	32.6	5.4	84495	4	US-09-797-906-3	Sequence 3, Appl
39	32.2	5.3	1255	4	US-09-227-357-52	Sequence 52, Appl
40	32	5.3	480	1	US-08-282-581-4	Sequence 4, Appl
41	32	5.3	480	1	US-08-550-544-4	Sequence 4, Appl
42	32	5.3	4084	3	US-08-866-340-1	Sequence 1, Appl
43	32	5.3	4460	4	US-09-103-875-4	Sequence 4, Appl
44	31.8	5.3	990	3	US-08-921-209-3	Sequence 3, Appl
45	31.8	5.3	990	3	US-09-411-763-3	Sequence 3, Appl

ALIGNMENTS

```
RESULT 1
US-09-228-986-63
Sequence 63, Application US/09228986
Patent No. 6359198
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Niels
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signalling
FILE REFERENCE: 11000/1020
CURRENT APPLICATION NUMBER: US/09/228,986
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 63
LENGTH: 1656
TYPE: DNA
ORGANISM: Pinus radiata
US-09-228-986-63

Query Match          9.6% Score 58.2; DB 4; Length 1656;
Best Local Similarity 55.6%; Pred. No. 1.1e-07;
Matches 133; Conservative 0; Mismatches 103; Indels 3; Gaps 1;

QY 72 TGTGATCGTGCATGATGATCCATTAACCTTAATTCATGAGAATCAAGCGA 131
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 TCTTGCTGCGAAGACACACAAATCAAGATTAATTCGAGGGTGTTCMAAGCC 610
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 132 TTGGGGGTATTCACAGACAGGGAATACGAGGAGGAGGAGGATATTCACCGTGCAG 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 611 TTATCTTACTGTGAGAGAGCTGAGATAGGAAAGTACAGTACATATTTCAAGCAGG 670
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 192 GCGGCTCATCTTGAACCTTAATCCATGATGAATGAAGATGCCAGAGGATGCTTT 251
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 671 GCAG---AACATATGATCTTGTGTATGACAAAGAGATGCTGTATGATGGCATG 727
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 252 CGACACTAGAGACGTAAGAGAAATGAGTAAGTCAATGATTTTGGGGGAGACTTCA 310
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 AGGCACACAGGACACTGATGATGAGTACGAGTACAGCCACTTGTTCACACTACAGCA 786
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
```

```

; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTzgtt-Fls
; US-08-232-463-14

Query Match
Best Local Similarity 7.5%; Score 54; DB 1; Length 7218;
Matches 33; Conservative 220; Mismatches 185; Indels 0; Gaps 0;

QY 1 GGCACGACGAGATGGAAGATGCGCAAAATCCATGGAGATATCGAGAAATAAAGAA 60
DB 1425 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1366
QY 61 GAACTAAAGTGTGATCGTATGATGATCCACTAAACCTTAATTCATGAGAAGAT 120
DB 1365 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1306
QY 121 CATCAAGGATGGGGGATTTTCACACAGCCGATACGCTGAGAGGAGCAATCAT 180
DB 1305 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1246
QY 181 CCACCGTACGCGGCTCATCTTTGACCTTATCTTAATGATTAAGAAATGCCGAG 240
DB 1245 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1186
QY 241 GGATGGTGTTCGACACTAAGACTAAGAAATGAGTAAGTCAATGATTTGGTGG 300
DB 1185 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1126
QY 301 GGTGACTGCTGCTGCAATGAAGAGAGCGAGGCTTTCATGAGAGCTGACTTAA 360
DB 1125 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1066
QY 361 CCATGCTTGGCAAAACGTTAACAGAGACAGATCATCCCTCTCAATTAACCACTCAT 420
DB 1065 TCGCAAGCTCCCTCGACCTGCGAGCCAGCTGGAATTAATCTGTGAGCGGTATGGCAAC 1006
QY 421 GGATGCTGATGATATA 438
```

```

DB 1005 GAAGCAAAAATAGTTATA 988

RESULT 3
US-09-078-294-3
; Sequence 3, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 80595
; TYPE: DNA
; ORGANISM: Nucleotide sequence of HC-contig
US-09-078-294-3

Query Match
Best Local Similarity 7.4%; Score 44.8; DB 4; Length 80595;
Matches 91; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 430 ATGATATATATTTATTTATTTAGAACACACATTAATAGCTCTAAGTGTATGTATCG 489
DB 1558 ATACATTAATATGATATATATGATATAGACATTAATATGATATATATATAGACATTA 15647
QY 490 ATAGATACCTGATGATGTTGTTTATAGATTTAGGTTCTTATTCGCTGATATATA 549
DB 15648 ATAAATATGATATATGATATATAGACATTAATATGATATATATATAGACATTA 15707
QY 550 TCATGTAAGTGTCTTCTTAAGCTTATTAATATTTAAATAGGCTTT 597
DB 15708 ATATGATATATATGATATATATAGACATTAATATGATATATATGATATAT 15755

RESULT 4
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

Query Match
Best Local Similarity 7.3%; Score 44; DB 4; Length 80246;
Matches 89; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 434 ATATATATTTATTTATTTAGCAACACATTAATACCTTAAGTGTATGTATGATAG 493
DB 15605 ATAAATATGATATATATGATATATAGACATTAATATGATATATATAGACATTA 15664
QY 494 ATATTCGATGTGTGTTTGAATTTAGGTTCTTATTCGCGCGATATATATCAT 553
DB 15665 ATATGATATATGATATATATAGACATTAATATGATATATATGATATATAGACATTAAT 15724
```


US-08-482-279-70	
Query Match	6.6%; Score 39.8; DB 2; Length 2551

```

; LENGTH: 2501 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```


;; FILING DATE: 18-NOV-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/930,649
;; FILING DATE: 14-AUG-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/889,020
;; FILING DATE: 26-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Campbell, Cathryn A.
;; REGISTRATION NUMBER: 31,815
;; REFERENCE/DOCKET NUMBER: P-UC 3003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 535-9001
;; TELEFAX: (619) 535-8949
;; INFORMATION FOR SEQ ID NO: 70:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2551 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-09-397-386-70

Query Match 6.6%; Score 39.8; DB 4; Length 2551;
Best Local Similarity 45.6%; Pred. No. 0.027; Indels 0; Gaps 0;
Matches 140; Conservative 0; Mismatches 167;

QY 298 TGGGCTGACTCACTGGCTGCAATGAAGAGGAGCGGCGCTTTCATGGAAGCTGGACT 357
DB 3 TGAGACCAACTGCTGATATACAGAAAGCAATTATGCTGTAATAGTAAGCTTGGT 62
QY 358 TAACATGCTTGGCAAAACCGTTAACCAAGGACAAATATCCCTCTCATTAACCACT 417
DB 63 TCTGACTGATAGATCTTACCCCTACAGAGATTCAAGTGGTGTGCTCCATTGAACAAT 122
QY 418 CATGATGCTGATGATGATATATTTATATTTATGAAACACACATTAATTAAGCTTAAGT 477
DB 123 AGTATATATGTTTAT 182
QY 478 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
DB 133 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 242
QY 538 CGTAT 597
DB 243 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 302
QY 598 CCTGAC 604
DB 303 CACTTAC 309

RESULT 11
US-09-031-563-26
;; Sequence 26, Application US/09031563A
;; Patent No. 6022708
;; GENERAL INFORMATION:
;; APPLICANT: Frederic de Sauvage
;; APPLICANT: Arnon Rosenthal
;; TITLE OF INVENTION: Fused
;; FILE REFERENCE: P1272
;; CURRENT APPLICATION NUMBER: US/09/031,563A
;; CURRENT FILING DATE: 1998-02-26
;; NUMBER OF SEQ ID NOS: 27
;; SEQ ID NO 26
;; LENGTH: 4586
;; TYPE: DNA
;; ORGANISM: Drosophila virilis
US-09-031-563-26

Query Match 6.6%; Score 39.8; DB 3; Length 4586;
Best Local Similarity 52.0%; Pred. No. 0.034;
Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 427 TTGATGAT 486
DB 4054 TTATGAT 4113
QY 487 TGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
DB 4114 TGTATGAT 4173
QY 547 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
DB 4174 TGTAT 4224

RESULT 12
US-09-078-294-4/C
;; Sequence 4, Application US/09078294
;; Patent No. 6265211
;; GENERAL INFORMATION:
;; APPLICANT: Choo, Kong-Hong Andy
;; APPLICANT: Du Sart, Desirée
;; APPLICANT: Cancilla, Michael R.
;; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
;; FILE REFERENCE: Davies Col
;; CURRENT APPLICATION NUMBER: US/09/078,294
;; CURRENT FILING DATE: 1998-05-13
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: Patent In Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 80246
;; TYPE: DNA
;; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

Query Match 6.4%; Score 38.6; DB 4; Length 80246;
Best Local Similarity 52.9%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 83; Conservative 0; Mismatches 74;

QY 434 AT 493
DB 15426 AT 15367
QY 494 ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
DB 15366 AT 15307
QY 554 GTAAGTGTGCTTAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
DB 15306 GTGTAT 15270

RESULT 13
US-09-078-294-3/C
;; Sequence 3, Application US/09078294
;; Patent No. 6265211
;; GENERAL INFORMATION:
;; APPLICANT: Choo, Kong-Hong Andy
;; APPLICANT: Du Sart, Desirée
;; APPLICANT: Cancilla, Michael R.
;; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
;; FILE REFERENCE: Davies Col
;; CURRENT APPLICATION NUMBER: US/09/078,294
;; CURRENT FILING DATE: 1998-05-13
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: Patent In Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 80595
;; TYPE: DNA
;; ORGANISM: Nucleotide sequence of HC-contig
US-09-078-294-3

Query Match 6.4%; Score 38.6; DB 4; Length 80595;
Best Local Similarity 52.9%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 83; Conservative 0; Mismatches 74;

QY	434	ATAATATTTTAAATATGGAACACACATATTAACGCTAAGGTTATATGATGATG	493
Db	15691	ATAATACAAATATTAATGCTATATACACATATTAACATATTAATGATATATACACATAT	156322
QY	494	ATATCGCATGCTGTGTTTGAATTTAGSGTCTTATAGCTCCGATCAATTAATCAT	553
Db	15691	ATACAAATTTATGCTATATACACATATTAACATATTTATGATATATATTTATGATATATAT	155722
QY	554	GTAAGTGTGCTTAAAGCTTTAAATATTTAAATA	590
Db	15571	GTCGTATATGATACATATATACATATACATACATACATATA	15535

RESULT 14
 US-08-629-643A-5
 Sequence 5, Application US/08629643A
 Patent No. 6025539
 GENERAL INFORMATION:
 APPLICANT: Lee, J. L.
 APPLICANT: Lee, N. A.
 TITLE OF INVENTION: IL-5 TRANSGENIC MOUSE
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Schweigman, Lundberg, Woessner & Kluth, P. A.
 STREET: P. O. Box 2938
 CITY: Minneapolis
 STATE: MN
 COUNTRY: USA
 ZIP: 55402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/629,643A
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/629,643
 FILING DATE: 09-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Viksnius, Ann S
 REGISTRATION NUMBER: 37,748
 REFERENCE/DOCKET NUMBER: 150.167W01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 612-359-3260
 TELEFAX: 612-359-3263
 TELEX:
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6727 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 650..3771
 OTHER INFORMATION: Join 650..730, 1560..1592,
 OTHER INFORMATION: 3468..3596, 3676..3771
 US-08-629-643A-5

[illegible]

Db 2623 GGTGATATTATATCATGCATGTCATGCATGATCATGCATGATCATATAT 2670

RESULT 15
US-09-280-799-1
; Sequence 1, Application US/09280799
; Patent No. 6136603
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Karras, James G.
; APPLICANT: McKay, Robert
; TITLE OF INVENTION: ANTISENSE MODULATION OF INTERLEUKIN-5 SIGNAL
; TITLE OF INVENTION: TRANSDUCTION
; FILE REFERENCE: ISPH-0340
; CURRENT APPLICATION NUMBER: US/09/280,799
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 6727
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-280-799-1

[illegible]

Search completed: November 7, 2002, 12:40:54
Job time : 293 secs


```

/db.xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-4294"
/clone.lib="Gm-cl019"
/tissue.type="Immature seed coats of greenhouse grown plants"
/lab.host="DH10B (Gibco BRL)"
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I; Site 2: Sal I. This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mg) of greenhouse grown plants. The library was prepared using the Life Technologies Superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."
BASE COUNT      133 a      66 c      88 g      89 t
ORIGIN

```

```

Query Match      12.5%; Score 75.4; DB 9; Length 376;
Best Local Similarity 34.6%; Pred. No. 9.5e-08;
Matches 195; Conservative 0; Mismatches 156; Indels 6; Gaps 2;

```

```

QY  62 AACCTAACGCTGTGATCGTCATGATGATCCAAACCTTATATTCATGAGAAGATC 121
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  26 AATTATACAGCACTGTGTAGTATGACATTAATCAACCAAGATTCATCAAAAGCTG 85
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  122 ATCAAAGCGATTGGGGGATTTTACACAGACCGAATAACGGTGAAGGAGCATGATC 181
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  86 TTGGAGAGTGTGGATGATGAATAATCAGAGAGTGGAAATGGCCAAAGACGATGACAT 145
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  182 CACCGTGAAGGCGGCGTCACTTTTGACCTTATCCATGATGAATAAGAAATGCCGAGAGC 241
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  146 CATGGCATGAC---MAAGATTGACCTGATTCATGACATGATGATGCCATATG 202
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  242 GATGCTGTTTGACAACTAAGAGTAAGAAATGGAAGTGAAGTGAATGATTTGGC 301
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  203 AATGGCATTTGAGCAACAAAGCAAGTCTGCTCAATGGGCGATGATGATGATTTGGT 262
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  302 GTGACTGCTGCTGACAAATGAAGAGGCGGCGCTTCATGAAAGCTGGACTTAAC 361
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  263 GTATCTACGCTGT---ACGAGACCAAAATACGAAATTTATGGAAGGCGGCTGAAT 319
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  362 CATGCTTGCGCAAAACCGTTAACCAAGACATCATCCCTCATTAACCAACTC 418
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  320 GACTACCATGAGAAACCGCTTGAACTTAAGCTTACCTTCTTGAATAGATC 376
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 2
BE822715/c      BE822715      568 bp      mRNA      linear      EST 24-MAY-2001
LOCUS          GmT00018B10H9 Gm-r1070 Glycine max cDNA clone Gm-r1070-7289 3'
DEFINITION     mRNA sequence.

```

```

ACCESSION      BE822715
VERSION        BE822715
KEYWORDS       EST.
SOURCE         soybean.
ORGANISM       Glycine max

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

```

```

REFERENCE
1 (bases 1 to 568)

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

COMMENT
Other ESTs: AM278862 corresponding to Gm-cl019-4294 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for

```

```

FEATURES
source
    Soybean (NSF 9872565)
    Lewin, H. A., Director, Keck Center for Comparative and Functional
    Genomics
    University of Illinois
    Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
    Tel: (217) 244-6147
    Fax: (217) 333-4582
    Email: l-vodkin@uiuc.edu
    This clone is available through: Genome Systems, Inc. 4633 World
    Parkway Circle St. Louis, Missouri 63134. For further information
    call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
    427-3324 or contact: clones@genomesystems.com or info@genome
    systems.com web site: www.genomesystems.com
    Seq primer: 5'-TTTCTTTTCTTTTCTTTTCTTTT(A/C/G)-3'.
    Location/Qualifiers
    1..568
    /organism="Glycine max"
    /db.xref="taxon:3847"
    /clone="Gm-r1070-7289"
    /clone.lib="Gm-r1070"
    /note="The library Gm-r1070 is a sequence-driven, reracked
    set of 9,216 clones selected from cDNA libraries from
    various tissues and stages of development of soybean that
    represent 2,639 sequences from immature cotyledons, 1,770
    from immature seed coats, 3,938 from flowers, and 869
    from young pods. The 5' ESTs of the source clones from
    the different libraries was used to select singletons, or
    a representative of each contig, which were reracked to
    form library Gm-r1070. The cDNA clones of the reracked
    Gm-r1070 library were then sequenced at the 3' end. The
    contig analysis to select unique genes was performed by
    the laboratory of Ernest Retzel, Center for Computational
    Genomics and Bioinformatics, University of Minnesota,
    http://www.cbc.umn.edu/ResearchProjects/soybean/index.html
    Reracking was performed by Genome Systems, St. Louis,
    http://www.genomesystems.com, and 3 sequencing by the
    Keck Center for Comparative and Functional Genomics,
    University of Illinois,
    http://www.life.uiuc.edu/biotech/keck.html. Note: The
    corresponding 5' EST from each clone in the Gm-r1070
    library is listed in the 'OTHER EST' field. The detailed
    information on the source library for each clone can also
    be obtained by referring to the Genome Systems clone ID of
    the original cDNA library that is also listed under
    'OTHER EST'."

```

```

BASE COUNT      163 a      110 c      95 g      179 t      21 others
ORIGIN

```

```

Query Match      10.4%; Score 63.2; DB 10; Length 568;
Best Local Similarity 53.0%; Pred. No. 7.8e-05;
Matches 175; Conservative 0; Mismatches 149; Indels 6; Gaps 2;

```

```

QY  90 ATCACTAAACCTTATTAATCATGAGAAGATCATCAAGAGATGGGGATTTTCAAGA 149
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  567 ATAAATATNNNNGAAAGATTCATCAAAAGCTGTGGAGATGTGGAATGAAATCAAC 508
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  150 CAGCGAATAACGGTGAAGAGGCACTAATCATCCACCTGACGGCGCTCATCTTTGACC 209
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  507 GAGTGAATAATGGCCAAAGACAGTGCATCTC---TTGCCATATGCAAAAGATTTGACC 451
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  210 TTATCTATGATGATTAAGAAATGCCCGAGAGGATGTGTTTGACAACTAAGAACTAA 269
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  450 TGATCTCATGAGACATGATGATGCCATCATGATGATGATGATGATGATGATGATG 391
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  270 GAGAAATGAAGTGAAGCAATGATTTGTTGGGATGATCTACGCTGACATGAAGAGG 329
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  390 GGTCAATGGGCAATGTGATGATGTGTGTGATATCAACCTGT---ACGAGACGAG 334
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  330 AGCGAGGCGCTTATGAGAGCTGACATCACTGCTTGGAACCAACCGTTAACCAAG 389
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  333 AATATGAAATTTATGAAACGGGACATGATGCTCATGAGAAACCTTGAACAAATT 274
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  390 ACAGATATATCCCTCTCATTAACCAACTCA 419
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```



```
/clone_lib="Hordeum vulgare 5-45 DAP spike EST library
HYCDNA0009 (5 to 45 DAP)"
/issue_type="5-45 DAP spike"
/lab_host="SOLR"
```

[note="Vector: lambdaZAP; Site1: EcoRI; Site2: XhoI; Plants were grown in the greenhouse at the University of California, Riverside (Fenton, St Close, TU Close). Whole spikes with awns trimmed were collected at 5, 10, 15, 20, 30 and 45 DAP (Fenton). Total RNA was prepared from each pool, equal quantities of all six RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give plus-script SK(-) cDNA phageids (XhoI) in the TU Close lab at the University of California, Riverside. Phageids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wind). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Willing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phased value 20 or above. For more details on library preparation and sequence analysis see

[illegible]

Query Match	8.3%	Score 50.2;	DB 10;	Length 732;
Best Local Similarity	49.3%;	Pred. No. 0.098;		
Matches 189; Conservative	0;	Mismatches 188;	Indels 6;	Gaps 2

58 GAGAACTAAACGTGTTCATCTGCATGATGATCCACTAAACCTTATATTCATGACAA 117

DG 33 GAGCAGACATCACGGGTCTCCCTAGTGTGCAGCGACCAGGAGAACCCACAGGGTTTGCGCGAAGGC 92

118 CATCATCAATAACCCCATTGCCCCCCCTTCCTACTCTATCTATCTAGCGGTCTTAAAGCGGGAAGGTA 176

[illegible][illegible][illegible][illegible][illegible]

295 TGTGTGGGCTGACCTTCACTGGGCTGACCAATTGAAAGAGAGAGCGCTATGGCAAGCCCTGG 354

273 CGTCGCCGTG---TCCAGGCACAGCCCTCCGTCGATGTCAGAGCCCTTCCATCCGCCCCCGCG 320

355 ACCTTAACCATTCGTTGGCCAAACCGTTAACCCBAAGGACACAGATCATTCCTTCATATAACCA 414

330 AGCCGACGACTTACGCCCAAGCCGTCACCAAGAGAGAGCTGGGCTACATCTCTCCAA 389

415 ACTCATGGATGCTTGATGGATAT 437

Db 390 GTTCGGGCTTGCGTAGCGGCCAT 412

RESULT 5
BM373897

LOCUS	BM373897	426 bp	mRNA	linear	EST 10-JAN-2002
DEFINITION	EBma03_SQ003_A24_R	IGF Barley	EBma03 library	Hordeum vulgare	cdna

clone EBma03_s0003_A24 5', mRNA sequence.
BM373897

VERSION BM3/3897.1 GI:18117287
KEYWORDS EST.

[illegible][illegible]

AUTHORS
 Realey, P., Liu, H., Caldwell, D., McCallum, N., Mule, S., Cardie, L.,
 Ramsay, L., Machary, G., Marshall, D.F.M. and Maughan, R.
 TITLE
 Development of Barley Transcriptome Resources
 JOURNAL
 Unpublished (2001)
 COMMENT
 Contact: Maughan R

Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731

Fax: 00 44 1382 562426
 Email: rwaughes@scri.sari.ac.uk
 All sequence has a Phred quality score of 20 or over
 Seq primer: M13 reverse.
 Location/Qualifiers

```
source
1. 426
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
```

```

/clone="Ema03_SQ003_A24"
/clone_lib="IGF Barley Ema03 Library"
/tissue_type="Maternal tissue"
/dev stage="8 days post anthesis"

```

```

/1ab_host="DH10B"
/notes="vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from maternal tissue dissected from developing

```

grains (8 days post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project "

Query Match	8.2%	Score 49.8	DA 10	length 426
BASE COUNT	114 a	66 c	108 g	138 t
ORIGIN	Genomic region chr1:100000000-100000000			

Best Local Similarity 54.6%; Pred. No. 0.12;
Matches 124; Conservative 0; Mismatches 97; Indels 6; Gaps 1;

[illegible][illegible][illegible]

Y 5/4 AAAACCGTAAACAGAGACAGAGATACCTCTCATTTAACCAATCAT 420
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 177 AAACCATGAGGAGGAGGTTCTATACCTTCATTCAGAGATCAT 223

LOCUS	BI509132	266 bp	mRNA	linear	EST 29-AUG-2001
RESULT 6	BI509132				

DEFINITION	Bee Brain Normalized/Subtracted Library, BB17 Apis mellifera cDNA clone BB17/0013B1E07 5', mRNA sequence.
ACCESSION	BB1509132
VERSION	BB1509132.1 GI:15359506

KEYWORDS EST. honeybee.
SOURCE Apis mellifera
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata
; Apoidea, Apidae, Apis.

	REFERENCE	X(bases 1 to 266)
AQ488569/c	TITLE	Whitfield,C.W., Soares,B.,Robertson,H.M., pardinas,J., Liu,L., Smoller,D. and Robinson,G.E.
DEFINITION	JOURNAL	An Expressed Sequence Tag Resource for Studies of Brain and Behavior In The Honey Bee Unpublished (2001)
ACCESSION	COMMENT	Contact: Gene E. Robinson Department of Entomology University of Illinois 505 S. Goodwin Ave., Urbana, IL 61801, USA Tel.: 217 265 0309 Fax: 217 244 3499 Email: genecro@life.uiuc.edu This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield. REPEAT IN THE SEQUENCE Simple repeat STRAND (+) ELEMENT (TA)n LOCATION [93..266]. PCR primers FORWARD: TAATAGCACTCATATTGAGC BACKWARD: ATTAACTTCATGAAG Insert Length: 266 Std Error: 0.00 Plate: BB17001JB10 row: E column: 07 Seq primer: ACCGTATAACAATTCCACACAGA High quality sequence stop: 266. Location/Qualifiers 1..266 /organism="Apis mellifera" /strain="mixed strains of European bees, predominantly A.m. ligustica" /db_xref="taxon:7460" /clone="BB17001JB10E07" /location.lib="Bee Brain Normalized/Subtracted Library; BB17"/sex="female" /tissue.type="brain" /dev_stage="adult worker honey bee" /lab_host="DH10B" /note="Organ: brain; Vector: pTn3-Pac; Site_1: EcoRI; Site_2: NotI; This Bpl7 cDNA library was generated by subtraction of the Bbl6 library with 4000 previously sequenced clones. The Bbl6 library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996). Genome Research 6(3): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."
BASE COUNT		113 a 16 c 26 g 11 t
ORIGIN		
Query Match		7.9%; Score 47.6; DB 10; Length 266;
Best Local Similarity		54.6%; Pred. No. 0.37; Indels 0; Gaps 0;
Matches		95; Conservative 0; Mismatches 79;
OY	417 TCATGGATCGTTGATGGATATATATTATTATATATGGAACACACTAATAADAGTGTAG	476
'Db	80 TGAATTCAGTATATATATATATATGAAAATATATATATATATATATATATATA	139
OY	477 TGCTATGATGCATAGATGACTGCCATGGTGCTGTTTAGAAATTAGAGTCCTTATCGT	536
Dd	140 CGGTATNGTA	199
OY	537 CCATGATATATATCATGTPAACGTGTTGCTTTAACCCTTAAAAATATTAAATA	590
Dd	200 TA	253

RESULT 7

LOCUS AQ488569 696 bp DNA linear GSS 24-APR-1999

DEFINITION RNCIT-1L-243M4.TJ RPCL-1l Homo sapiens genomic clone RNCIT-1L-243M4,

DATA DESCRIPTION

ACCESSION AQ488569

VERSION	AAQ88569.1	GI:4674443
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 (bases 1 to 696)	
JOURNAL	Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter	
COMMENT	J.C. Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building Unpublished (1997) Other_GSSs: RPCI-11-243N4.TV Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbeet@ig.org Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pletier de Jong (pletier@edj.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html . Seq. primer: SP6 Class: BAC ends.	
FEATURES	Location/Qualifiers	
Source	1..696	
	/organism="Homo sapiens"	
	/db_xref="GDB:7593243"	
	/db_xref="taxon:9606"	
	/clone="RPCI-11-243N4"	
	/clone_lib="RPCI-11"	
	/sex="Male"	
	/cell_type="Lymphocytes"	
	/note="Vector: pPhoc3.6; Site.1: EcoRI; Site.2: EcoRI; RPCI11 Human Male BAC Library"	
BASE COUNT	236 a 93 c 106 g 260 t	1 others
ORIGIN		
Query Match	7.8% Score 47.4; DB 12; Length 696;	
Best Local Similarity	55.9%; Pred. No. 0.45; 71; Indels 0; Gaps 0;	
Matches	90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	430 ATGGATATATATTTTATATTATGAAGAACACATATAAAGCTTAAGCTGTATGATGC	489
Db		
581 ATAATTTTATTTTATATATATATAATAACATATATATAAAATTCGTGTATATATGT		522
OY	490 ATAGATCTTGCACTGTGTGTTTAGAATTTAGAGGCTCTTATCGCCGTGATATATA	549
Db		
521 AATAATCATATATGTGTATATACACATATAACTGTGTATACGTGTATACACATAT		462
OY	550 TCATGTAGTGTGCTTACGCTTATATAATTTAATA	590
Db		
461 ACGTGTGTATACGTGTATATACAACATATATGTATATA		421
RESULT_B		
BM368320	442 bp MRNA linear EST 10-JAN-2000	
LOCUS	BM368320	
DEFINITION	Ebed01_SQ002_N02.R IGF barley Ebed01 library Hordeum vulgare cdna	
ACCESSION	BM368320	
VERSION	BM368320.1 GI:18111710	
KEYWORDS	EST.	
SOURCE	barley.	
ORGANISM	Hordeum vulgare	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae	
	; Triticeae; Hordeum.	
	1 (bases 1 to 442)	

AUTHORS
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
Ramsay, L., Machray, G., Marshall, D.F.M. and Maugh, R.
TITLE
Development of Barley Transcriptome Resources
JOURNAL
Unpublished (2001)
COMMENT
Contact: Maugh R
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rmaugh@scri.sari.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse

FEATURES
Source
Location/Qualifiers
1..442
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBed01.S0002.N02"
/clone_lib="IGF Barley EBed01 library"
/tissue_type="Endosperm"
/dev_stage="6 days post anthesis"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from endosperm tissue dissected from developing
grains (6 days post anthesis) in glasshouse grown barley
plants. Developed as part of the barley transcriptome
resources of BBSRC/SEERAD funded cereal IGF (Investigating
Gene function) project."

BASE COUNT
113 a 89 c 120 g 120 t
ORIGIN

Query Match
Best Local Similarity 7.8%; Score 47.2; DB 10; Length 442;
Matches 119; Conservative 0; Mismatches 93; Indels 6; Gaps 1;

QY 203 TTGACCTTATCTATGATTAAGAAATGCCGAGAGGATGTTGACACATAAG 262
|||||
DB 218 TTGTGATTTGTTTGTGATTAAGACATGCCCATCATGATGTCGCCGAGCAGTTG 277
QY 263 AAGCTAAGAGAAATGAGATGAATCAATGTTGGGTGACTTCAGCGCTGACAT 322
|||||
DB 278 AAGATCCCTGCTATGAGGAGAGTGATGTAAGATTTTGGGGTTCTGCGCATGATCAT 337
QY 323 GAGAGGAGCGGAGGCTTTTCATGGAAGCTGACTTAACCATGCTTGGCAAAACGTTA 382
|||||
DB 338 GCATATGA-----GGCATTCATGATGCTGCTGATGCTTGGCCCAACCAATG 391
QY 383 ACCAAGGACAGATCATCCCTCATTAACCAACTCAT 420
|||||
DB 392 AGGAGGAGGCTTCTCATCTATCATCAAGAGATCAT 429

RESULT 9
LOCUS BM415113 1745 bp mRNA linear EST 28-JUN-2002
DEFINITION Op0184 Mixed Stage EST's from Globodera pallida, the potato cyst
nematode Globodera pallida cDNA, mRNA sequence.
ACCESSION BM415113
VERSION BM415113.1 GI:18381463
KEYWORDS EST.
SOURCE Globodera pallida.
ORGANISM Globodera pallida.
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heterodermidae; Heterodermineae; Globodera.
REFERENCE
AUTHORS Heer, J., Sosinski, B., Pokrzywa, R.M., Wary, A. and Opperman, C.
TITLE Mixed Stage EST's from Globodera pallida, the potato cyst nematode
JOURNAL Unpublished (2001)
COMMENT Contact: Opperman, C
Center for the Biology of Nematode Parasitism
NC State University, IACR-Rothamsted

Campus Box 7616, Raleigh, NC 27695, USA
Tel: 919.515.6599
Fax: 919.515.9500
Email: warthog@unity.ncsu.edu
Gill-11PCN_R_H05_G711-11_R_044.ab1.
Location/Qualifiers
1..1745
/organism="Globodera pallida"
/db_xref="taxon:36090"
/clone_lib="Mixed Stage EST's from Globodera pallida, the
potato cyst nematode"
/note="Vector: lambda GT11; This is a collaborative effort
between IACR-Rothamsted and North Carolina State
University. The library was constructed from mixed stage
G. pallida in lambda GT11 by Paul Burroughs,
IACR-Rothamsted."

BASE COUNT 439 a 472 c 262 g 547 t 25 others
ORIGIN

Query Match
Best Local Similarity 7.6%; Score 45.8; DB 10; Length 1745;
Matches 89; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 430 ATGATATATATTTTATATATGGAACACACATTAACGCTAGTGTATGATGCC 489
|||||
DB 901 AT 842
QY 490 ATAGATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
|||||
DB 841 ATATATATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
QY 550 TCATGATAGTGTGCTTAACTTAAGCTTAAATATATTAATA 590
|||||
DB 781 ATATCTATATCTA 741

RESULT 10
LOCUS CNS030D0 1024 bp DNA linear GSS 15-MAY-2000
DEFINITION Tetradon nigroviridis genome survey sequence 17 end of clone
184002 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL222093.1 GI:7880912
VERSION AL222093.1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetradontiformes;
Tetradontidae; Tetradon.
1 (bases 1 to 1024)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
Unpublished
2 (bases 1 to 1024)
Roest-Crollius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C.,
Bernot, A., Fizes, C., Wincker, P., Bottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 1024)
Genoscope.

TITLE
JOURNAL Direct Submission
REFERENCE Submitted (12-APR-2000) to the EMBL/Genbank/DBP databases
AUTHORS This sequence is a single read and was generated as part of a large
genome. For more information, please take a look at
COMMENT http://www.genoscope.cns.fr/Tetradon.

```

FEATURES
  source      Location/Qualifiers
              1..1024
              /organism="Drosophila melanogaster"
              /db_xref="taxon:99883"
              /clone="184002"
              /clone_11b="g"
              /note="Genoscope sequence ID : C0AG184BH01UP1-end : 17"
BASE COUNT   346 a      70 c      174 g      333 t      101 others
ORIGIN
Query Match   7.5%; Score 45.4; DB 12; Length 1024;
Best Local Similarity 46.0%; Pred. No. 1.4;
Matches 74; Conservative 18; Mismatches 69; Indels 0; Gaps 0;

QY 430 ATGATATATATATATATATATGGAACACACATATAAGCTAGTGTATGATGTC 489
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 WTAATATATATATATATATATATATATATATATATATATATATATATAT 577
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 ATGATATCTGCATGTGTGTATGAATTTAGGCTCTTATGCTCGCATATATATA 549
    : ||| ||| : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 TTTTAAATTTTWTAAATATTTTAAATATATATATATATATATATATATAT 517
    : ||| ||| : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 TCATGATGTTGCTTAAGCTATAAATATATATATAA 590
    : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 ATAAATATATATATATATATATATATATATATATA 476
    : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
LOCUS      1225 bp      DNA      linear      GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
            BACR15C18 of Drosophila library from Drosophila melanogaster (fruit
            fly) genomic survey sequence.
ACCESSION  AL106171.1 GI:5620504
VERSION     AL106171.1
KEYWORDS   GSS.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 1225)
AUTHORS   Direct Submission
            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBelobac11.
FEATURES
  source      Location/Qualifiers
              1..1225
              /organism="Drosophila melanogaster"
              /plasmid="pBelobac11"
              /db_xref="taxon:7227"
              /clone_11b="DrosBAC"
              /clone="BACN15C18"
              /note="end : Sp6"
BASE COUNT   266 a      128 c      38 g      368 t      425 others
ORIGIN
Query Match   7.5%; Score 45.4; DB 12; Length 1225;
Best Local Similarity 28.0%; Pred. No. 1.4;
Matches 118; Conservative 93; Mismatches 210; Indels 0; Gaps 0;

QY 177 TCATCAGCGTGAAGGCGCTGATCTTATGCTTATGCTAATGATAGAAATGCGG 236
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 TCCCTCCYCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 738
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 237 AGAGGATGGTGTTCGACACTAGAAAGCTAGAGAAATGGAAGTGAATGATGATG 296
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 KAAAMAAATACCKBKKNKKKKKKMMADAAKAAAMAAHMTTATTAHAAYNA 798
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 297 TTGGGGTGACTTCAGTGGCTGACAAATGAAGAGGAGGAGGCTTCATGGAAGCTGAC 356
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 AAATTTAAATCCKAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAM 858
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 357 TTACCATGTCCTTGCAAAAACCGTTAACCAAGCAAGATACCCCTCATTACCAAC 415
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 AANRABDDKAAAAAATATATATATATATATATATATATATATATATATATATAT 918
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 417 TCATGATGCTGATGATATATATATATATATATATATATATATATATATATAT 476
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 919 ATAAATATATATATATATATATATATATATATATATATATATATATATATAT 978
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 477 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 536
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 TTTTAAATTTTAAATATATATATATATATATATATATATATATATATATAT 1038
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 CCGTGAATATATATATATATATATATATATATATATATATATATATATATAT 596
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1039 TTTTAAATTTTAAATATATATATATATATATATATATATATATATATATAT 1098
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 597 T 597
Db 1099 T 1099

RESULT 12
LOCUS      525 bp      DNA      linear      GSS 23-AUG-2000
DEFINITION gen_UMB001_050_A08F Glycine max genomic clones containing resistance
            gene analogs from UMN Soybean BAC Library (PECSBAC4 EcORI) Glycine
            max genomic, DNA sequence.
ACCESSION  A2301735
VERSION     A2301735
KEYWORDS   GSS.
SOURCE     soybean.
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eudicots I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE  1 (bases 1 to 525)
AUTHORS   Pennella,S., Larson,K., Mudge,J., Damesh,D., Denny,R.,
            Foster-Hartnett,D. and Young,N.D.
            BAC end sequences from soybean BACs containing resistance gene
            analogs (RGAs)
            Unpublished (2000)
            Contact: Young Nevlin D
            Department of Plant Pathology
            University of Minnesota
            495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul MN 55108, USA
            Tel: 612 625 2225
            Fax: 612 625 9728
            Email: nevin@tc.umn.edu
            Sequence on contig Gm_NBSDB8_ctg_b near unmapped duplicate of
            NBSDB8.
            Seq primer: M13F
            Ends: BAC ends.
FEATURES
  source      Location/Qualifiers
              1..525
              /organism="Glycine max"
              /cultivar="Paribault"
              /db_xref="taxon:3847"
              /clone_11b="Glycine max genomic clones containing
              resistance gene analogs from UMN Soybean BAC Library
              (PECSBAC4 EcORI)"
              /tissue_type="cotyledon leaves"
              /dev_stage="cotyledon stage"
              /note="Vector: pECSBAC4; Site_1: Eco RI; The UMN BAC

```

`/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"`
`/note="Vector: PWD42nv; Purified genomic DNA from M.`
`musculus C57BL/6J (male) was obtained from the Jackson`

Email: szhao@igr.org
 Clones are derived from the mouse BAC library RPEC-23. For BAC
 library availability, please contact Pieter de Jong

